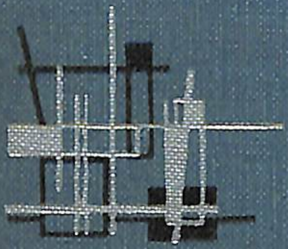
FOURTH EDITION



INDUSTRIAL PSYCHOLOGY

JOSEPH TIFFIN
ERNEST J. McCORMICK

© 1942, 1947, 1952, 1958, BY

PRENTICE-HALL, INC.
ENGLEWOOD CLIFFS, N.J.

ALL RIGHTS RESERVED. NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY MIMEOGRAPH OR ANY OTHER MEANS, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHERS.

LIBRARY OF CONGRESS
CATALOG CARD NO.: 58-11417

150.13
TIF

Bureau Ednl. rch
E
Dated 21-7.59
Accs No 6417

PRINTED IN THE UNITED STATES OF AMERICA
46238

To
Mary and Emily

PREFACE

During recent years there has been a marked increase in the application of psychological methods and techniques to the solution of many human problems in industry and business. This increased reliance upon psychology in industry and business probably is, in large part, the consequence of two major factors. In the first place, the importance of human beings has received wider recognition in industry and business. Secondly, during these years many developments have taken place in the field of industrial psychology that have extended the potential applications of this science to a wide variety of human problems in industry and business. The applications of industrial psychology that are covered in this book fall primarily into two broad areas. One of these deals with the *personnel* function, including such things as the selection and placement of employees, employee training, merit rating, job evaluation, and motivation. The other area is concerned more with adapting work to the capacities and limitations of human beings; this area is sometimes called *human engineering*. The usefulness of psychology in these functions has been clearly demonstrated in many different types of situations.

Although substantial gains have been made in the application of psychology to many of the problems of industry and business, there remain many challenging problems to be solved. One such challenge is that of developing methods for creating work situations within which employees may gain greater satisfactions; another is that of reducing the frictions in labor-management relations.

Thus, industrial psychology, by its very nature, is dynamic. In view of continual developments in the field, it has been the intention of the authors to include, in this fourth edition, a considerable amount of the research that has been completed since the previous edition.

In addition to the inclusion here of much of the recent research in industrial psychology, there are also some changes in organization and in emphasis from the earlier editions. In particular, considerably greater emphasis has been placed on employee motivation and morale, and on human engineering. On the other hand, there is somewhat less emphasis on employment tests, and, in particular, on vision and vision tests. These shifts in emphasis tend to reflect corresponding shifts in the focus of industrial psychology.

The ever-growing body of research in industrial psychology is the result of the efforts of many individuals. This group includes a number of former graduate students of Industrial Psychology at Purdue University. Much of the work of these people is incorporated in this book. We also take this opportunity to express our appreciation to those individuals whose published work has been made available to us. Acknowledgment is also made to the many publishers for granting permission to reproduce figures, tables, and other materials.

In addition, we would like to make a special acknowledgment to several of our colleagues at Purdue University who have in various ways contributed to the development of this book. These men include Professors E. J. Asher, C. H. Lawshe, N. C. Kephart, B. J. Winer, and W. V. Owen.

Though the first three editions of this book were written by the first-named author, the present edition is the result of joint authorship. It is hoped that the somewhat different backgrounds and interests of the authors have resulted in a comprehensive and balanced coverage of the material.

JOSEPH TIFFIN
ERNEST J. MCCORMICK

CONTENTS

1. **INTRODUCTION** 1

2. **THE SIGNIFICANCE OF INDIVIDUAL DIFFERENCES IN INDUSTRY** 11

Individual Differences of Job Applicants, 15
The Effects of Training upon Individual Differences, 20
Individual Differences in Job Performance, 24
Individual differences in productivity. Individual differences in merit ratings. Individual differences in other aspects of performance. Stability of job performance. Importance of individual differences.
The Bases of Individual Differences, 30
Criteria of Job Success, 32
Types of criteria. Selection of the criterion. Contamination of the criterion. Reliability of the criterion. Combining criteria. Importance of the criterion.

3. **THE INTERVIEW AND RELATED METHODS OF EMPLOYMENT** 39

Employment-Office Procedures, 39
Methods of Determining Personnel Specifications, 41
Personnel specifications set by judgments. Personnel specifications set by job analysis. Personnel specifications set by statistical analysis. Consolidating personnel specifications. Uniqueness of job specifications.
Information About the Applicant, 52
The interview. The application form. References. Personnel tests.

4. GENERAL PRINCIPLES OF PERSONNEL TESTING **75**

General Concepts, 76

Testing supplemental to other employment procedures. Extent of use of employment tests. Selection or placement as the function of tests. Psychological tests are not infallible. Types of tests. Definition of terms.

Testing the Tests—The Experimental Approach, 81

Present-employee method of testing the test. Follow-up method of testing the test. Comparison of the two methods of testing the test. Other methods of test validation.

Factors Determining the Functional Value of Personnel Tests, 91

Validity. Reliability. Relationship between reliability and validity. The selection ratio. Percentage of present employees considered satisfactory.

Combining Tests into a Battery, 105

Adequate Training of Testers, 108

5. APTITUDE TESTS **110**

Kinds of Aptitudes, 110

Mental Ability Tests, 110

The nature of mental abilities. Typical mental ability tests. Examples of mental ability tests in use.

Mechanical Ability and Related Tests, 122

The nature of mechanical ability. Group tests of mechanical ability. Individual tests of mechanical ability. Examples of mechanical ability tests in use.

Psychomotor Tests, 134

The nature of psychomotor abilities. Typical psychomotor tests. Examples of psychomotor tests in use.

Vision Tests, 145

The nature of visual skills. Examples of vision tests. Examples of vision tests in use.

Tests of Special Aptitudes, 156

Clerical aptitude tests. Perception tests.

Validity of Aptitude Tests for Various Occupations, 157

6. TESTS OF PERSONALITY AND INTEREST 161

Tests of Personality, 163

The Bernreuter Personality Inventory. The Humm-Wadsworth Temperament Scale. The Guilford Series of Personality Tests. The Minnesota Multiphasic Personality Inventory. The Thurstone Temperament Schedule. Gordon Personal Profile and Personal Inventory. The Rorschach Projective Test.

Interest Tests, 171

Strong Vocational Interest Blank. Kuder Preference Record. Cardall Primary Business Interests Test.

Limitations of Personality and Interest Tests, 174

Faking of personality and interest tests. Fakability of forced-choice tests.

The Use of Personality and Interest Tests, 178

The validation of personality and interest tests. Validity of personality tests for employment. Other uses of personality and interest tests.

7. ACHIEVEMENT TESTS 185

Validity of achievement tests.

Oral Achievement Tests, 186

Validation of oral trade questions. Factors affecting the validity of trade questions. Industrial uses of oral achievement tests.

Job Sample Tests, 190

A miniature punch press. A performance test for forklift operators. The job sample in clerical work.

Written Achievement Tests, 195

The Purdue Vocational Tests.

Uses of Achievement Tests, 199

Employment. Transfer and promotion of employees. Training. Measurement of vocational achievement.

8. INDUSTRIAL MERIT RATING 204

The Growth of Merit Ratings, 204

Purposes of Merit Rating, 205

INDUSTRIAL MERIT RATING (continued)

Administrative uses of merit ratings. Employee self-improvement.

Considerations of Merit vs. Seniority, 208

Merit Rating Systems, 209

Rating scales. Employee comparison systems. Check list rating systems. Critical incident technique.

Some Dangers of Merit Rating, 221

The halo effect. Constant error. Influence of other factors on ratings. Giving rating information to employees.

Adjusting Ratings for Undesired Influences, 235

Adjusting for differences in means. Adjusting for differences by standard scores.

The "Weighting" of Traits, 237

Reliability and Validity of Merit Ratings, 242

Reliability of ratings. Validity of ratings. Combining ratings from different raters. Effect of rater's memory on ratings.

A Merit-Rating Case Study, 247

Suggestions for Improving a Merit-Rating Program, 247

9. TRAINING OF INDUSTRIAL EMPLOYEES 251

Determining Training Needs, 252

Check lists of training needs. Judgments of training needs. Job analysis for training purposes.

Characteristics of Learning, 261

The learning curve. The plateau.

Principles of Learning, 265

Knowledge of results. Motivation. Attention to specific features of the job. Transfer of training. Negative transfer or interference. Distribution of practice.

Training Methods and Aids, 272

The Measurement of Training, 275

Production. Production time required to do the job. Number of operators meeting production standard. Increase in learning rate. Decrease in breakage or use of consumable supplies. Reduction in number of acci-

TRAINING OF INDUSTRIAL EMPLOYEES (continued)

dents. Reduction in absenteeism. Reduction of labor turnover. Measurement of information.

Types of Training, 287

On-the-job training. Apprentice training. Supervisory and management training.

10. THE MEASUREMENT OF ATTITUDES AND MORALE 302

The Meaning of Terms, 302

Attitudes. Opinions. Morale.

Measurement of Attitudes and Opinions, 304

Attitude scales. Opinion surveys. "Theme" analysis of employee communications. Interviews with employees. The exit interview.

Job Satisfaction, 317

Job satisfaction in general. What people want from their jobs.

11. ATTITUDES AND HUMAN BEHAVIOR 327

Employee Attitudes and Behavior, 327

Employee attitudes and job performance. Employee attitudes and absenteeism, tardiness, and tenure. Comments.

Understanding Human Behavior, 334

Causation in human behavior. Human needs and incentives. Frustration.

12. HUMAN RELATIONS 340

Supervision, 340

Supervision and employee performance. Selection of supervisors. Training of supervisors. Self-improvement of supervisors. Characteristics of supervisory behavior.

Communications, 346

What is meant by communications? Methods of communication. Effects of communications. Nature of communicative ability. Readability of written communications.

HUMAN RELATIONS (continued)

Employee Participation, 360

Types of employee participation. Examples of employee participation. Discussion.

Group Unity, 368

Differences in group behavior. The development of group unity.

Conclusion, 372

13. INDUSTRIAL INSPECTION 373

Measuring the Accuracy of Inspectors, 374

Accuracy in tin-plate inspection. Accuracy in the use of precision instruments.

Placement of Inspectors, 384

Placement of tin-plate inspectors. Placement of electrical inspectors. Placement of bottle selectors.

Methods of Inspection, 388

Methods of inspecting tin-plate. The use of examples as standards. Layout of work area. Method of display.

Training of Inspectors, 395

Training of tin-plate inspectors.

Summary, 397

14. ACCIDENTS AND SAFETY 398

General Considerations, 398

Psychological vs. physical conditions as causes of accidents. Lost-time accidents, home cases, and first-aid cases. Accident proneness. Accident proneness vs. the job in relation to accidents.

External Factors Related to Accident Proneness, 409

Psychological Factors Related to Accident Proneness, 412

Vision. Age and plant service. Emotional factors. Mental ability. Ratio between perceptual and muscular speed. Other psychological factors. Identifying factors related to accident proneness.

How to Reduce Accident Proneness, 427

ACCIDENTS AND SAFETY (continued)

Placement of employees. Safeguarding and training inexperienced employees. Personal protection. Immediate attention to injury.

15. WORKING CONDITIONS 433

Criteria in Evaluating Working Conditions, 433
Physiological criteria. Psychological criteria. Work output. Other criteria. Relationships among criteria. Discussion of criteria.
Illumination, 441
Illumination and visual acuity. Illumination and brightness contrast. Optimum illumination levels.
Atmospheric Conditions, 445
The heat-exchange processes. Conditions that affect heat exchange. Effects of atmospheric conditions on people. Comfort zones.
Noise, 450
Effects of noise on work performance. Effects of noise on hearing. Noise control.
Music, 454
Hours of Work, 455
Rest Pauses, 457

16. HUMAN ENGINEERING AND WORK METHODS 460

Man-Machine Systems, 461
Information-Receiving Processes, 462
Visual displays.
Decision-Making Processes, 465
Nature of the decisions to be made. Manner of presenting information.
Action-Taking Processes, 469
Blind-positioning movements. Size of handwheels. Workspace.
Methods Analysis, 473
Motion economy as related to the human body. Measuring effects of methods.

17. WAGES AND JOB EVALUATION 482

Factors Affecting General (Community and Industry) Wage Levels, 483

Supply and demand. Governmental wage controls. Civic relations of the company. Cost of living. Regional and industrial differentials in workers' earnings. Strength of organized labor. Cost of production.

Factors Affecting Wage Rates Within a Plant, 486

Incentive vs. hourly paid jobs. Nonfinancial incentives. Discrimination against women and other groups. Company policy. Supply and demand. Collective bargaining agreement. Job evaluation.

General Procedures in Installing and Operating a Job Evaluation Program, 489

The job evaluation committee. Obtaining the cooperation of supervision. Job descriptions for job evaluation. Evaluation of jobs. Converting evaluations to money values. Provisions for evaluation of new jobs. Getting jobs in line after the system is installed.

Job Evaluation Systems, 492

Scaling of jobs by comparing entire jobs (ranking method). Scaling of jobs by comparing components of each job (factor comparison method). Comparing entire job with a predetermined scale (classification method). Comparing job elements of each job with a predetermined point scale (point method). Job description check list.

Converting Job Evaluation Results to a Wage Scale, 497

Weighting of Items in Job Evaluation, 500

Statistical Studies of Job Evaluation Plans, 501

Reliability of Job Evaluation, 509

APPENDIXES

A. ELEMENTARY DESCRIPTIVE STATISTICS 513

Descriptive vs. Sampling Statistics, 513

Graphic Representation of Data, 514

ELEMENTARY DESCRIPTIVE STATISTICS (continued)

The frequency distribution and polygon. The normal distribution.

Measures of Central Tendency, 519

The arithmetic mean. The median. The mode. When to use the mean, median, and mode.

Measures of Variability, 524

The standard deviation.

Comparable Scores, 527

Percentiles, 532

Correlation, 534

Rank-order correlation. The product-moment coefficient.

B. TAYLOR–RUSSELL TABLES 548

C. PUBLISHERS OF REPRESENTATIVE PERSONNEL TESTS 554

Intelligence tests. Clerical tests. Industrial vision tests. Stenographic and typing skills. Dexterity and manipulative tests. Mechanical tests. Personality tests. Interest tests. Trade tests. General.

Publishers, 557

INDEX 559

INDUSTRIAL PSYCHOLOGY

1

INTRODUCTION

Psychology is a new field of study. Though there is much evidence of interest in studying the nature of man as far back as recorded history—or even legendary history—the very beginning of systematic study in this field dates only from 1879 with the founding of the first psychological laboratory by Wundt in Leipzig. Even since that date, progress in the field has been halting and there have been periods when decades passed with no apparent—and perhaps little real—progress at all. Although every branch of psychology is therefore young, *industrial* psychology is probably the youngest branch of all. The first book on industrial psychology was probably one by Münsterberg,[1] published in 1913. And yet modern industry has accepted and is using a great deal of the work of the psychologist. Perhaps one reason for this acceptance is that the "rules of the game" in American industry have undergone drastic changes during the past quarter century and some of the techniques of psychology seem to offer help in adapting to the new order.

A few contrasts between the industrial personnel situation in the relatively recent past and the situation today may serve to highlight the kinds of changes that have occurred. Not so long ago industrial management tended to regard labor predominantly as a cost of production. Every effort was made to keep this cost, like any other, as low as possible. In attaining this objective *on a short-time basis*, many policies were followed that management now realizes are not economical in the long run. Wages were set by rule of thumb. There was little or no formal training for either operators or supervisors. New employees often were simply told to "go to work" and "ask the man on the next bench if you have any questions." There were no provisions for medical or health insurance. Pension plans for retirement at some specified age were almost nonexistent.

---

[1] Hugo Münsterberg, *The Psychology of Industrial Efficiency* (Houghton Mifflin Company, 1913).

New employees frequently were selected by "methods" that resulted in no better than chance placement of the right man on the right job. The former practice of "hiring at the gate," for example, gave no assurance whatever that the men hired would be satisfactory employees. Interviewing for employment was conducted by people who had little or no background for such work. And the interviewers seldom had available the factual type of information about the applicants that is now usually provided for use in assessing the potentiality of applicants in terms of the requirements of the jobs for which they are being considered.

In quite a few companies, employment methods were used that bordered on mystic, occult practices. These methods included graphology (judgment of personal traits from handwriting), physiognomy (judgment from the shape of the face and head), phrenology (judgment from the shape and contour of the skull), and even astrology (judging a man's future activities from the position of heavenly bodies at the time of his birth).

As late as 1941, a questionnaire sent by Long and Tiffin to twelve major companies that were said to be using graphology brought in enthusiastic endorsements of this pseudoscience from nine of the companies.[2] The vice-president of an important railroad wrote as follows:

> What graphology does tell, however, which a mental test does not, is the temperament of the individual. It is easy to tell whether a man is an introvert or an extrovert; his attitude toward other people; whether he is a research man or a salesman; a man fond of himself; and in the hands of an expert it tells many other things concerning mental and physical characteristics.

The vice-president of a furnace manufacturing company wrote as follows:

> . . . this filled-in application form is submitted to a graphologist—more for character analysis than for any other purpose—and our experience over the past four years indicates that the percentage of misses on the part of the graphologist are less than 10%.

---

[2] W. F. Long and Joseph Tiffin, "A Note on the Use of Graphology by Industry," *Journal of Applied Psychology*, 1941, 25, 469–471.

Executives of seven other companies wrote in a similar enthusiastic manner about how well personality and character could be evaluated from an applicant's handwriting.

A number of psychologists have conducted well-controlled experiments to determine whether there is any real relationship between handwriting and personality characteristics. Several of these studies have been summarized by Allport and Vernon.[3] The relationships found between various handwriting characteristics and measurable personality tests were no higher than could be accounted for by chance.

Similar rigid scientific studies have been made of physiognomy and phrenology.[4] These investigations have clearly shown that such methods have no validity whatever, and that judgments about people based on these "signs" are no better than can be accounted for by chance.

Such ineffective hiring methods could be tolerated by management in the days when men could be laid off (or terminated or fired) without notice and without a systematic plan. The decision about who should be terminated rested entirely with the company, usually with the immediate supervisor. When an employee was found to be unsatisfactory—a day or a year after he had been employed—he was immediately dismissed. Needless to say, many men were laid off for various reasons that they could not fully understand. The growing strength of organized labor and a parallel growth of social consciousness on the part of employers have resulted in a definite change in policy with regard to layoff. Seniority now carries great weight in many companies. In such companies, after an employee has served a probationary period, usually sixty to ninety days, he begins to accumulate seniority. When layoffs must be made because of a decline in business, the last men hired usually are the first to be laid off. When this is the rule to be followed, it is obvious that management should exercise every precaution to be certain that an employee is really satisfactory before he begins to accumulate seniority. And entirely aside from the dissatisfaction of the working force when men are arbitrarily dismissed, this practice results in excessive turnover that is undesirable because it is expensive. A number of investigations have shown that the average cost of losing an employee, hiring a new one, and bringing the new one up to a level of standard efficiency is at least $200. For financial reasons alone, if for no other, management is now making every effort to keep turnover at a minimum.

---

[3] Gordon Allport and Philip E. Vernon, *Studies in Expressive Movement* (The Macmillan Co., 1933).

[4] G. E. Cleeton and F. B. Knight, "Validity of Character Judgment Based on External Criteria," *Journal of Applied Psychology*, 1924, **8**, 215–231.

Whenever men work with or for each other, disputes occasionally arise. In the industrial situation this must be expected. A way must therefore be found to settle such disputes. Until relatively recent times the rule that had been practiced for generations was still almost universally followed—the supervisor or foreman (i.e. management) was right and the employee was wrong. Decisions made by the foreman were not always entirely fair to the employee and were even at times detrimental to the best interests of the company. This situation has been changed in many companies (and always where a union has been certified) by the establishment of a formal grievance procedure. Such a procedure provides for several steps that must be taken to reach a settlement of the grievance; and, if prior settlement is not reached, recourse is finally had to arbitration. An outside arbitrator is called upon to examine all the facts pertaining to the dispute and to render a decision. Most agreements between a company and a union provide that the decision of the arbitrator will be final and binding on both parties.

A few years ago supervisors and foremen were almost invariably men promoted from the ranks because of outstanding job performance at the nonsupervisory level. Such men had long been accustomed to a driving type of leadership. This was expected by management and the men had seen foremen prosper for years by being drivers rather than leaders. The successful foreman today is a different type of man. Emphasis is placed upon the human relations aspects of the foreman's job, and supervisory training courses on such topics as "How to Understand Employees" have received wide support by management. A questionnaire survey of thirty-six large corporations by Stagner revealed that 83 per cent operated a training program to show supervisors how to deal with men.[5] As Toedt has pointed out, the modern philosophy of personnel administration involves obtaining cooperation and understanding between management and the workers all down the line.[6] The human and social aspects of the job as well as technical skills must be considered to obtain effective and economical performance.

---

[5] R. Stagner, "Attitudes of Corporate Executives Regarding Psychological Methods in Personnel Work," *American Psychologist*, 1946, 1, 540–541.

Many more such changes in personnel policies have occurred during recent years and still more changes are now under way. Most managements today have accepted the point of view that a dissatisfied and disgruntled working force often increases the personnel phase of production costs completely out of proportion to material and equipment costs. The changes have paralleled the growing strength and recognition of organized labor during this period. Prior to the National Labor Relations Act (the Wagner Act), which became law in 1935, a management did not need to accept or even to recognize a union. Furthermore, an employee could be discharged for membership in a union, for suspicion of membership, or even for making remarks favorable to a union. All companies had been by no means so intolerant as they could have been under the previous law, or lack of law, but enough had been antagonistic to have alienated large segments of organized labor. The Wagner Act gave employees the right to elect the union of their choice and required the company to bargain collectively with that union to reach a written agreement covering wages, working conditions, and various phases of company personnel policy. The changes during these years in the attitudes of management toward employees have been accompanied by many important developments in personnel practices. Granting that there are still many out-dated procedures still in existence, there is a growing trend toward a more rational and objective approach to the solution of personnel problems in business and industry.

In order to make more adequate decisions regarding personnel problems, however, it is necessary to have pertinent factual information. In many other aspects of business and industry, management would not *think* of making decisions *without* having certain important *facts* at hand. Through the techniques and procedures of industrial psychology, it is now possible to carry out systematic investigations that provide management with pertinent information regarding personnel problems. One of the common functions of an industrial psychologist is that of conducting personnel research—research directed toward the collection, analysis, and organization of information about people in the work situation that can be used effectively by management in dealing with such problems.

---

[6] Theodore A. Toedt, "The Modern Philosophy of Personnel Administration," *Personnel Journal*, 1953, **32**, 178–180.

In carrying out any type of research, it is necessary to measure, or quantify, the variables that are involved. In the physical sciences measurement usually can be fairly accurate and not too difficult (e.g., measuring the expansion of gas under different temperatures, or the tensile strength of a piece of steel). The measurement of *human* variables, however, frequently is quite difficult and usually it is not as precise as physical measurements. Psychologists, however, have developed methods of measuring many of the human variables with sufficient accuracy to serve useful purposes. The measurement devices involve such things as personnel tests, rating procedures, and attitude questionnaires. The point we are leading up to is that industrial psychology has now developed to a stage where it is making a distinct contribution to the solution of many personnel problems in industry through the process of personnel research.

Aside from performing such research activities, industrial psychologists also serve in two other basic capacities. One of these is consultation on personnel problems. By reason of his knowledge of human behavior (especially in the work situation), the psychologist frequently is able to give valuable assistance to management on such matters as training employees how to do certain jobs or how to get acceptance of a new safety program.

The other activity of the industrial psychologist might be called an *installation* function. Psychologists frequently are asked to help "install" some new feature of a personnel program, such as one dealing with personnel testing, merit rating, or job evaluation.

In general, the various applications of psychology in industry fall into three areas. The first of these embraces the *personnel activities* of management. The previous discussion was concerned with these, and it should be added that this book will deal primarily with this area of industrial psychology. Some of the particular topics in this area include personnel selection and placement, training, attitude measurement, employee motivation, merit rating, safety, communications, and job evaluation.

A second area, one of growing importance, is what is coming to be known as *human engineering*. This includes the design of equipment and work environment in terms of the capabilities and limitations of people. The purposes of this branch of industrial psychology are, of course, the improvement of the effectiveness of work and making work easier for people to perform. We will discuss this area briefly in Chapters 15 and 16.

A third area of work in the broad field of psychology applied to industry is that associated with *consumer behavior*. There are various aspects of this area including market research, advertising, selling, and the study of consumer preferences. While this is an important area of psychology, it is beyond the scope of the present book and, therefore, will not be covered.

The primary orientation of this book will be concerned with the personnel activities of management. The acceptance by management of the industrial psychologists' work in this area is shown by a table of results obtained from the questionnaire survey by Stagner referred to previously.[7] The questionnaire of eleven items was mailed with a covering letter to the presidents of 70 of America's larger corporations. The corporations selected were in the fields of electrical, chemical, mechanical, and specialty manufacturing, food and beverage processing, mining, and retail trades. The questionnaires were returned by 36 of the 70 companies. Table 1.1 gives the percentage of respondents who answered "yes" to each of the eleven questions and also the percentage of respondents who answered "yes" to the

TABLE 1.1

Responses of Thirty-six Large Corporations on Psychological Methods in Personnel Work

	Does Your Company Have? (% of yes replies)	*Do You Consider this Desirable?* (% of yes replies)
1. A unified personnel office	72	75
2. Tests of intelligence for applicants	56	70
3. Tests of performance for applicants	58	64
4. Preference in employment to college men	58	44
5. Periodic check on the morale of employees by attitude tests	14	47
6. Professionally trained psychologists in the Personnel Department	30	53
7. Training program for new employees	78	83
8. Training program to show supervisors how to deal with men	83	81
9. Formal system for the expression of grievances	83	75
10. Employee counseling system	42	53
11. Employee merit rating system	67	70

---

[7] Stagner, *op. cit.*

question: "Do you consider this desirable?" It will be noted that more than 50 per cent of the companies were using intelligence tests for applicants, tests of performance for applicants, a training program for new employees, a training program to show supervisors how to deal with men, a formal system for the expression of grievances, and an employee merit rating system. It will be noted further in Table 1.1 that still higher percentages of the respondents to the questionnaire considered desirable the practice of intelligence tests for applicants, performance tests for applicants, a training program for new employees, and a merit rating plan. Every one of these topics has been studied extensively by psychologists. Although only 30 per cent of the responding companies reported the employment of professionally trained psychologists, 53 per cent of the respondents felt that such professionally trained people would be desirable.

It is recognized, of course, that the questionnaire returns summarized in Table 1.1 are based on a small sample of corporations, but the total number of people employed by these 36 organizations is roughly 815,000. The results thus show quite clearly that the larger corporations are making definite use of psychology in their personnel operations. The results from smaller companies would probably not be so encouraging because the personnel department of a small company would not be able to support extensive professional work of this type.

Even though psychological methods have been accepted by many of the larger corporations, the growth of psychology in industry has undoubtedly been retarded because management has often been unwilling to spend very much on procedures and methods that do not offer promise of immediate practical results. In obtaining initial support from industry it has, therefore, often been necessary to frame the psychological program in a way that does not involve very much, if any, long-range research. Industry ordinarily buys specific things such as a personnel selection program, a training program, or an attitude survey. There are many things in these and other areas of applied psychology on which the final answer or even the best methods of approach are not yet known. There is a real need for additional basic research, but such research is undertaken only spasmodically except by a few companies with a long history of successful use of psychological methods. Such companies, outstanding among which are Proctor & Gamble and General Motors, conduct a great deal of basic research that has revealed new techniques and results that are being used by many other companies. These companies, however, are the exception rather than the rule.

One matter that will be emphasized repeatedly in this book is the importance of measuring results. Once a change in a personnel program has been made, a follow-up study should be undertaken to compare what follows with the situation before the change was made. Several barometers of the effectiveness of a personnel program have been mentioned by Blai.[8] These include labor turnover, disciplinary actions, grievances filed, absenteeism, and reasons for voluntary terminations given in exit interviews. To these should be added at least two other barometers—namely, the morale of employees and the relative number of applicants for employment. Not all of these indicators are always available, or suitable in particular situations, but they suggest the *kind* of things that should be used to measure the effect of any change in a personnel program.

One retardant to the acceptance of psychology in industry, as Gorsuch has pointed out,[9] is that psychologists sometimes have difficulty selling their services in industry because there is confusion as to what they should or can contribute, and because they often lack orientation to the industrial situation. Psychologists are engaged in many different types of activities in industry today; employers need to know that different duties require different qualifications. And psychologists entering industry should recognize their own limitations. The student of industrial psychology should make a determined effort to relate the subject-matter of his course to the practical problems of the work situation.

Another retardant to the development of new psychological methods in industry arises from the fact that management is very cautious about using or even trying methods or procedures that might possibly cause some kind of industrial relations trouble. The very growth of the strength of organized labor has made management very cautious in trying methods that have not actually been proved in practice. But enough has actually been accomplished in industry to place psychological methods upon a firm footing today.

---

[8] Boris Blai, Jr., "How Effective Is Your Personnel Program?" *Personnel Journal*, 1951, **30**, 98–101.

[9] John H. Gorsuch, "Industrial Psychology's Growing Pains," *Personnel*, 1952, **29**, 153–157.

This book will illustrate the potential contributions of industrial psychology to the solution of problems involving people in work situations. In particular, it will emphasize the importance of a systematic, quantitative approach to such problems, directed toward providing organized, factual information that can be used in working out solutions of these problems.

2

THE SIGNIFICANCE OF INDIVIDUAL DIFFERENCES IN INDUSTRY

Supervisors and managerial personnel are keenly aware of the differences in the work performance and behavior of individuals. Even when several people are working on identical machines it will usually be found that there are differences in their productivity.

These differences in work performance and behavior arise because of differences in the personal characteristics of individuals. These differences run the gamut of human traits, including differences in physical build, in appearance, in intelligence, in aptitudes, in personality, in interests, in motivations, in physical skills, and in other traits. People also differ in their training, education, and acquired knowledges and job skills.

Individual differences are of primary consequence to business and industry as they affect the work performance or on-the-job behavior of employees. A job may be done well by one employee, fairly well by another, and very poorly by a third. The job may be a source of great personal satisfaction to one, a monotonous and boring task to another, and entirely beyond the capacity of a third. Modern industry is becoming increasingly aware of the importance of placing on every job an individual who is not only able to *do* the job well but who, in addition, is temperamentally adapted to the job in question. The success of personnel placement depends upon placing individuals on jobs that match their capacities and upon giving them adequate and specific training to do their jobs. If the job is too difficult or the training has been inadequate, the result is confusion, low production, and possible injury either to the operator or the machine. If the job is too easy, the result may be boredom, mind-wandering, and daydreaming, with the dissatisfaction that so often accompanies these activities.

An individual is best adapted and is usually most satisfied—and this applies whether he is at work, at home, or on his vacation—when he has found an outlet for whatever energy, drive, and ability he may possess. If his job calls for abilities that he does not have and cannot develop, he continually experiences the despair of failure. If, on the other hand, his job calls for only a fractional part of his ability, he is likely to develop other means of self-expression, which, at their best, may take the form of daydreaming or an unduly critical attitude and, at their worst, may become a definite and serious mental illness or disease. Energy not demanded by the job is usually released into some other channel. Too often the release of this "extra energy" is in a direction that not only fails to benefit either the employee or his employer but also, on the contrary, is actually detrimental to the interest of one or both of these parties.

The inadequate job performance of some employees has led managements of some companies to take various corrective steps such as the improvement of selection and placement programs. The improvement of such programs, however, requires the consideration, at the time of employment, of those personal characteristics that are related to job performance.

The implied relationship between personal characteristics and job performance has been demonstrated by many investigations. As early as 1923, for example, Bills[1] reported a significant relationship between intelligence and "difficulty of work" being done by clerical workers; in general, the more intelligent persons were on the most difficult jobs. She noted that after two and a half years the relationship was much more pronounced. This she interpreted to mean that, with time, there is a definite tendency for employees to shift, or to be shifted, into jobs of difficulty levels that match their respective abilities. In another early investigation, Pond and Bills[2] found that labor turnover can be markedly reduced by a careful placement of employees in jobs of a difficulty commensurate with the ability of those employees. Low-ability employees—as revealed by mental tests at the time of employment—showed only half the turnover on certain simple jobs as did high-ability employees assigned to the same jobs. On more difficult jobs, exactly the opposite situation prevailed: the low-ability employees showed the largest turnover and the high-ability employees the least. Standardized mental tests[3] reveal that on jobs of a repetitive, routine nature it is not uncommon to find negative correlations between productivity and mental ability. Such studies indicate the *practical* importance of placing employees on jobs that match their abilities.

---

[1] M. A. Bills, "Relation of Mental Alertness Test Scores to Positions and Permanency in Company," *Journal of Applied Psychology*, 1923, **7**, 154–156.

[2] Millicent Pond and M. A. Bills, "Intelligence and Clerical Jobs. Two Studies of Relation of Test Score to Job Held," *Personnel Journal*, 1933, **12**, 41–56.

Personnel managers, in hiring tradesmen, generally try to make every effort to determine in advance the degree of skill that the applicant possesses. But often employers do not so clearly recognize the fact that differences in capacity for machine operation and other jobs that are ordinarily considered as unskilled or semiskilled are just as great and just as important as differences in skill among tradesmen. The significance to industry of individual differences among employees is far more important than that of differences in skill already developed or of differences that can be detected in an interview. The concept of individual differences is concerned with basic differences in capacity which are of importance in every phase of industrial personnel placement.

Recent social legislation and the inclusion of some form of seniority clause in almost every union contract make the consideration of individual differences at the time of hiring more important than the placing of men and women on jobs that are neither above nor below their capacity to succeed. There was a time, not so long ago, when any employee—new or old—could be dismissed whenever his supervisor felt that the employee's services were no longer advantageous to the company. Often employees were tried out on several jobs over a period of months and were then dismissed as unsuited for any of the jobs available. Under modern legislative enactments and union-management agreements, it is practically impossible for industry to follow this procedure today. Unemployment insurance premiums are paid at least partly by the employer, and the schedule of premiums is so adjusted in many states that the amount that an employer pays is proportional to the labor turnover of the company. Union contracts often make it difficult to discharge an employee after he has been employed for a specified period of time.

---

[3] Joseph Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, 1939, 23, 240–263. See also J. Tiffin and C. H. Lawshe, Jr., "The Adaptability Test: A Fifteen Minute Mental Alertness Test for Use in Personnel Allocation," *Journal of Applied Psychology*, 1943, 27, 152–163.

At least two other factors operate to prevent the arbitrary dismissal of employees. One of these is based upon financial considerations, for the cost of training an employee is lost when that employee is dismissed. Training costs are no small item in the operation of a modern plant—a matter that will be discussed in some detail in Chapter 9. It should be remembered also that to the training cost, as such, must be added the wages paid to the employee while he is learning the job. Since March 1, 1956, a federal minimum wage law has required that every employee be paid at least $1 per hour if the company is involved in interstate commerce. This regulation applies to almost all manufacturing firms. By using every precaution to make sure that training is given only to those employees who are likely to stay on the job, at least until the investment in their training has been realized, and by making every effort to avoid the dismissal of these employees for *any avoidable reason*, sizable savings in production costs can almost invariably be made.

Another less mercenary, but very real, reason for a growing reluctance on the part of management to dismiss employees arbitrarily is the increasing sense of social obligation felt by a very large number of employers. At one time management regarded labor simply as a cost of production, and *only* as a cost of production. Today management recognizes that employees not only *make* the product, but also *buy* it; and that employees who leave dissatisfied may, and frequently do, spread discontent among other employees or future employees, a situation that works to the disadvantage of both producer and consumer. Satisfactory employer-employee relations are the cornerstone upon which a modern industrial organization must be built. For such reasons it has become more and more important for industry to evaluate the suitability of prospective employees *before hiring*. Differences in the suitability of applicants is a branch of the psychology of individual differences that, fortunately, has been quite thoroughly studied. To show the direct and practical application of these studies in the field of personnel administration will be a major purpose of this book.

In order to evaluate the potential suitability of applicants for a specific job, it is necessary to determine the characteristics that differentiate between better and poor employees on the job. In effect this requires a study of a sample of workers on the job in order to ascertain on what characteristics the better and poorer workers differ. If such differences are found, it is then possible to select, as new employees, those who have characteristics similar to the present better employees on the job. The procedures for making such studies will be described more in detail in later chapters. These procedures are mentioned here briefly, however, to emphasize the point that for such practical purposes it is necessary to measure, or identify, those individual differences that are related to job performance.

Particular reference has been made here to individual differences as related to job performance. The differences among individuals also have other implications in connection with their working relationships, as in the character of their relationships with fellow-employees and with superiors and subordinates, and in the satisfactions they gain from their work situations.

With reference to our major interest in individual differences as related to work performance, there are three contexts in which they need to be considered, as follows: (1) the differences among applicants that can be measured or identified at the time of application; (2) the effects of training upon individual differences; and (3) the differences in actual work performance that exist after completion of training. These will be discussed separately.

INDIVIDUAL DIFFERENCES OF JOB APPLICANTS

The person responsible for employment of applicants needs to take into account whatever pertinent information is available to him about the individual applicants in evaluating their potential suitability for various jobs. The "information" typically is available from the application form, from the interview, and from records of any employment tests the applicant has taken. Such sources of information will reveal differences among applicants on such characteristics or attributes as personal appearance, personality, verbal facility, height, weight, education, work experience, and aptitudes or achievement as measured by tests.

Examples of the differences that are typically found among applicants come from the records of a large manufacturing company that was considering applicants for the job of machine shop apprentice. Several tests were given to the applicants, these including the Otis Self Administering Test of Mental Ability, a standard intelligence test described briefly on p. 112, and the Technical Information in Machine Shop Test of the Purdue Vocational Series.[4]

The distribution of scores on these two tests are shown in Figs. 2.1 and 2.2. The height of each curve shows the per cent of individuals at the various score levels. To the extent that these tests measure

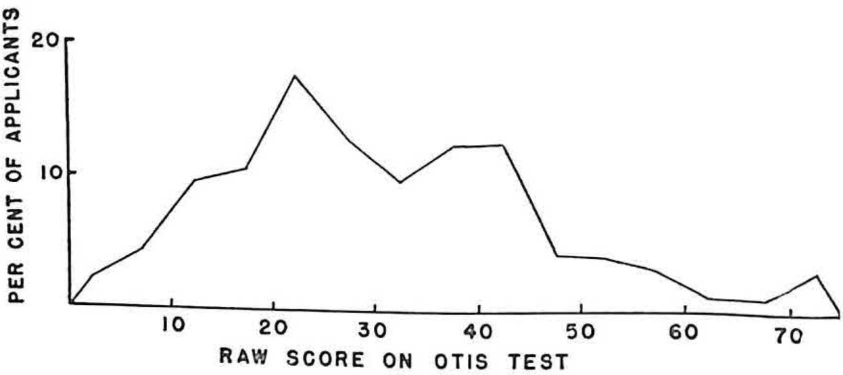


FIG. 2.1. Distribution of scores on Otis Self-Administering Test of Mental Ability (Higher Exam, Form A) of 112 applicants for machine shop apprenticeships.

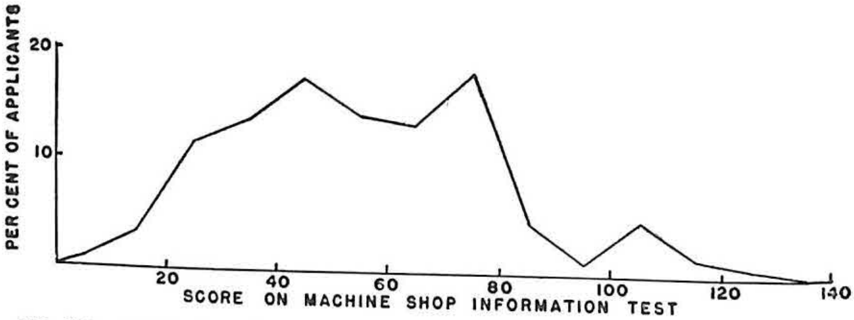


FIG. 2.2. Distribution of scores of 112 applicants on 130-item multiple-choice test covering technical information in machine shop practice.

---

[4] H. F. Owen, C. C. Stevason, H. G. McComb, and C. D. Hume, *Technical Information Test for Machinists and Machine Operators* (Chicago, Science Research Associates, 1942). This test is described briefly on p. 196. The original form of this test, on which Fig. 2.2 is based, contained 130 items. The revised form BH, distributed by Science Research Associates, contains 132 items.

variables that are related to success in apprentice training, it can be seen that the potentialities of individuals for such training vary a great deal. It might be added that all of the applicants either had taken high-school vocational courses or had had industrial experience as machinists' helpers. In spite of such opportunities to learn the technical details of the craft, the test on technical information in machine shop practices revealed enormous differences among the applicants in actual qualifications for the apprenticeship openings.

The above examples illustrate the range of individual differences of a group of applicants in terms of scores on an intelligence test and on a technical knowledge test. Let us take as another example the differences among people in finger dexterity, which is a skill that is necessary for successful performance on many typical production jobs. Such differences are illustrated in Fig. 2.3, which shows frequency distributions of four different groups[5] on the O'Connor test of finger dexterity.[6] This test consists of a 100-hole pegboard which is filled as rapidly as possible (three pins to a hole) from the pins located in a shallow tray. The greater one's finger dexterity, the more rapidly he will be able to fill the board and the smaller will be his score, which is simply the time in minutes required to fill the board. In Fig. 2.3 the left side of the scale represents the small-time values or good scores, and the right side the long-time values or poor scores. The upper curve, or *A*, in Fig. 2.3 shows the distribution of finger dexterity—as measured by this test—among a very large number of randomly selected persons. The remaining distributions in this figure (B, C, and D) show how finger dexterity is distributed among three groups of employees engaged in work for which finger dexterity is presumably an important qualification. A marked similarity among the four distributions is revealed. Operators on the three industrial jobs are little better in finger dexterity than are randomly selected persons, which means that persons equally qualified for this job would have been obtained if the employees had been selected by lot from a telephone directory.

The question will naturally be raised: What of it? How do we know that the finger-dexterity pegboard is a job qualification test for

---

[5] Tiffin and Greenly, *op. cit.*

[6] M. Hines and J. O'Connor, "A Measure of Finger Dexterity," *Personnel Journal*, 1926, 4, 379–382.

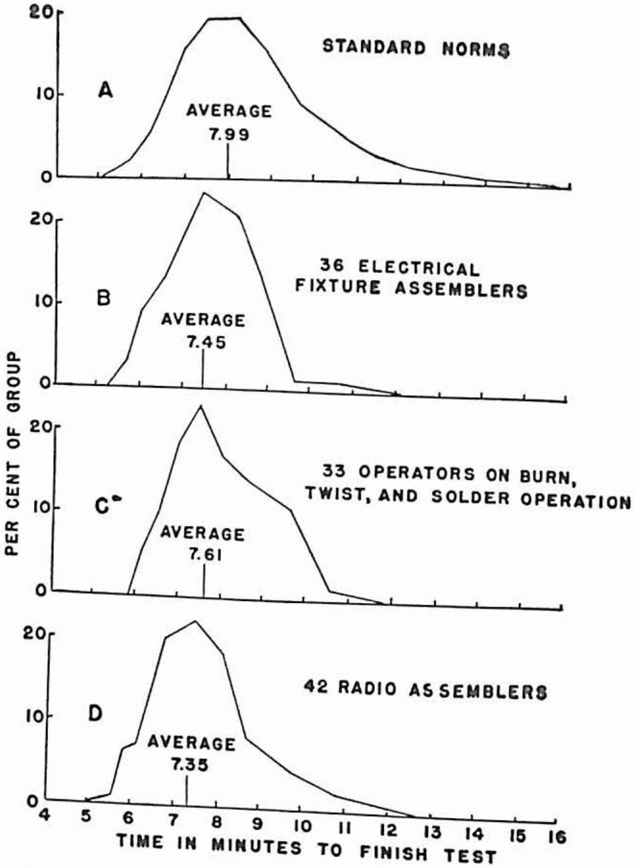


FIG. 2.3. Frequency distributions of scores on O'Connor Finger Dexterity Test of random group of subjects (top) and three groups of industrial employees engaged in rapid finger work.

these jobs? Are the persons who test high in finger dexterity better employees than those who test low?

Detailed answers to these questions must be postponed at this point (they will be given in Chapter 5); but we may definitely say here that the test in question does measure something that is necessary for efficient production on the jobs under consideration, and we may add that the employees testing high on the finger-dexterity test are, in general, those employees who are also most productive at their work.

As an example of still another area of individual differences let us take the matter of interests. The activities that interest one person may be completely boring to another. The strength of a person's interest in his job can have a major influence on his performance. Because of the importance of interests in vocational success, various interest tests have been developed, including The Kuder Preference Record.[7] This test has various subscores, each one reflecting the relative strength of interest in a separate type of work activity. The profiles of two individuals (X and Y) are presented in Fig. 2.4 to illustrate differences in interests. This figure shows, for each subtest, or interest area, the percentile rank of each person. The percentile rank is the per cent of individuals whose scores are exceeded by the person in question. The marked differences in such profiles certainly support the adage that one man's meat is another man's poison.

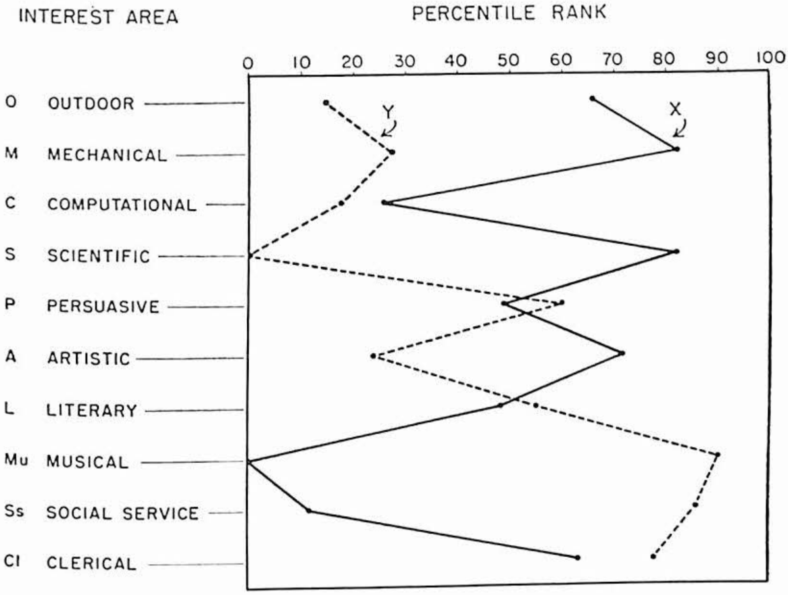


FIG. 2.4. Profiles of two individuals on the Kuder Preference Record. The percentile rank for each interest area is the per cent of persons whose scores are exceeded by the individual in question.

An employment program can be considered to be a success if it results in the selection and placement of individuals who, in general, perform effectively on their jobs. The crux of such a program, however, depends, in the first place, upon knowing what kinds of human characteristics are required for successful performance on each of the various jobs; and it depends, in the second place, upon being able to identify or measure such characteristics at the time of employment. Thus, the job of the employment manager is very largely one of sizing up the characteristics of individuals and evaluating them in relation to the requirements of job openings.

---

[7] The Kuder Preference Record is available from Science Research Associates, 57 West Grand Avenue, Chicago 10, Illinois.

Occasional mention has been made of the "measurement" or "identification" of individual differences. In employment-office operations some such differences are a matter of objective record; this is typically the case with such information as biographical data on the application form. For some characteristics, scores on employment tests may be available. Occasionally rating scales are used by interviewers for recording their judgments on such characteristics as appearance, speaking ability, or personality traits. In many instances the "information" about individuals is not a matter of record, but is in the form of subjective impressions in the minds of those who have interviewed the applicants. In general, the development of systematic employment procedures depends in large part upon the availability of reasonably objective, quantified information about applicants. Although subjective impressions of applicants must of necessity be used in employment activities, frequently they can be reduced to relatively quantitative terms, such as by the use of rating scales.

THE EFFECTS OF TRAINING UPON INDIVIDUAL DIFFERENCES

With new employees some type of formal or informal training almost invariably takes place. Many companies also have training programs for experienced employees, to help them to maintain their present skills or to develop new skills, or to qualify employees for higher-level positions. When employees undergo training, it is important to consider the effect of such training upon the individual differences that exist among the employees at the beginning of the training. Three possible conclusions might be considered: training may increase the differences, it may decrease them, or it may leave them unaffected.

Examination of experimental studies in this field reveals what seems to be much contradictory evidence. In certain kinds of situations it seems clear that training tends to increase whatever differences in ability may exist at the beginning of the training. A group of persons placed upon an entirely new task of a complicated nature are not likely to differ much among themselves in their ability to handle that task at the beginning. Their skill is likely to be quite uniform and, it might be added, uniformly low, until appropriate training has given those with capacity to do the task an opportunity to forge ahead. The operation of any complicated industrial production machine will furnish an example of this principle.

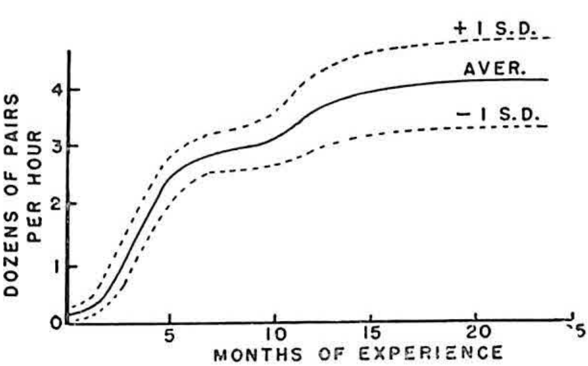


FIG. 2.5. Relation between experience and production in looping hosiery, and the effect of experience on individual differences in productivity. Results based on 199 loopers of varying experience.

In Fig. 2.5 the heavy black line shows the average production of hosiery loopers in dozens of pairs per hour plotted against the amount of experience on the job. This curve is really the learning curve for this operation, even though it was obtained by taking the average production of loopers of varying amounts of experience rather than by following the improvement in skill of a given group of loopers as experience increased. The dotted lines in Fig. 2.5 indicate a standard deviation[8] above and below the average production for

---

[8] The standard deviation, a measure of variability or spread of measurements, is explained in Appendix A, p. 525.

150.13
TIF

BUREAU OF EDNL. & PSYL. RESEARCH
1419
LIBRARY
21.3.59

the varying amounts of experience. In effect, the dotted lines show the range of performance of most of the loopers, specifically about the middle two-thirds of them. It is clear that the variability is greater among the loopers with considerable experience on the job than among those who have been recently hired or who are in the early stages of learning the operation.

It is customary in comparing two measures of variability, computed from two averages that are different, to make the comparison in terms of the coefficient of variation rather than in terms of the respective standard deviations. The coefficient of variation (CV)[9] is given by:

$$CV = \frac{100\ S.D.}{M}$$

where S.D. is the standard deviation of the measures
M is the mean of the measures

The coefficient of variation was found to be 100 for the employees with one month or less of experience and 225 for employees with experience of two years or more.

Here, then, is a specific case in which training *increases* the magnitude of individual differences with performance upon a given job. The operators are much more nearly like one another in the early stages of their training than in the intermediate or later stages.

One should not infer from the foregoing illustration, however, that the effect of training is always to increase the magnitude of individual differences. For example, Hartmann,[10] using four tests of manual ability, found that three months of training resulted in definitely more improvement among those who initially tested low than among those who initially tested high. The ability of everyone was improved as a result of the training, but, in this experiment, the training also tended to decrease the difference between the good and the poor workers.

It would be possible to cite numerous other experiments that support the apparently contradictory conclusions of one or the other of the investigations cited above. A study of the experiments in question, however, suggests a possible explanation for the seeming discrepancy in results. Wherever the effect of training has been to increase the magnitude of individual differences, a study of the task involved usually reveals it to be fairly complicated. By "complicated" we mean that it is one in which the average individual does not reach his maximum level of performance without a fairly long learning period. Wherever training results in a decrease in individual differences, the task is usually found to be a rather simple one, that is, one upon which the average individual shows little improvement after a relatively short practice period. In the light of these facts we may venture the tentative generalization that training tends to increase individual differences in proportion to the complexity of the task in question. Although a few minor exceptions to this principle may come to mind, it seems to be at least provisionally acceptable as a working hypothesis.

---

[9] J. P. Guilford, *Fundamental Statistics in Psychology and Education* (3rd ed.; New York, McGraw-Hill Book Company, Inc., 1956), 101.

[10] G. W. Hartmann, "Initial Performance as a Basis for Predicting Ultimate Achievement," *School and Society*, 1929, **29**, 495–496.

Whatever may be the effect of training on the magnitude of individual differences, it seems clear that training seldom changes the relative standing of individuals in their ability to perform any given task. This conclusion is in accord with the findings of Bishop,[11] Peterson,[12] Hartmann,[13] Viteles,[14] and others. No industry, therefore, should expect its training program to bring all employees up to the same high level of efficient performance unless the employees have initially been selected in terms of their capacity to reach that high level. And since perfect predictors of job success have not yet been devised, individual differences in employee performance will be found even if every effort is made to select only those who are most likely to succeed. Proper selection, however, can result in the employment of individuals who, collectively, will be *less* variable, and generally *better,* in job performance than would be the case with an unselected group of employees.

---

[11] M. K. Bishop, "A Study of Individual Differences in Learning," *Psychological Review*, 1925, 32, 34–53.

[12] J. Peterson, "The Effects of Practice on Individual Differences," *Twenty-Seventh Yearbook* of the National Society for the Study of Education, Part II (Bloomington, Ill., Public School Publishing Company, 1928), 212.

[13] Hartmann, *op. cit.*

[14] Morris S. Viteles, *Industrial Psychology* (New York, W. W. Norton & Company, Inc. 1932), 107.

INDIVIDUAL DIFFERENCES IN JOB PERFORMANCE

The tendency for relative individual differences to be maintained throughout training forecasts the conclusion that, in general, those individuals whose performance in training is above average, also will be those whose continuing performance on the job will be above average. Granting that employees on a given job do differ in their performance, one might then ask such questions as these: Are the differences large enough to be of practical importance? To what extent do they tend to be maintained over a continuing period? By what means can differences in employee performance be quantified for comparison purposes? One can answer such questions most satisfactorily by considering and interpreting statistical data available from various pertinent industrial situations. Various types of indexes of job performance can be used to illustrate the range of individual differences in job performance. Probably the most typical examples are those based on productivity of employees on production-types of jobs.

Individual Differences in Productivity. One example of differences in productivity is shown in Fig. 2.6. This is a frequency distribution showing differences in productivity among 36 electrical-fixture assemblers who were engaged in identical jobs. An incentive method of wage payment was in use and the best operator was earning slightly more than twice the average hourly wage of the poorest

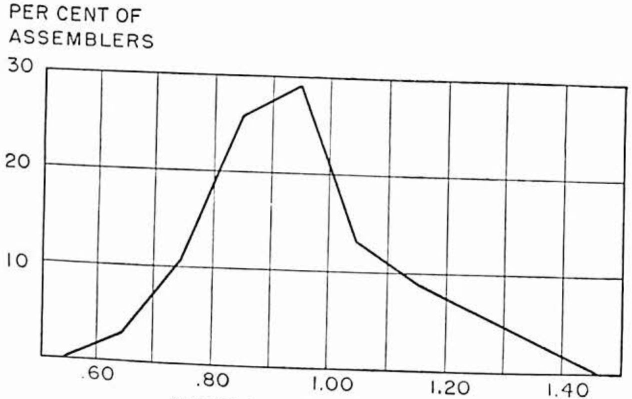


FIG. 2.6. Distribution of quantity of production among 36 electrical-fixture assemblers.

operator. The overhead to the company was identical for all the employees in this department, including the best and the poorest. Both used identical work layouts and assembling equipment. Both used identical amounts of space, heat, light, and other overhead expenses. But the best employee, in delivering more than twice the production of the poorest employee, was therefore using only half as much overhead and capital expense *per unit of production* as the poorest employee. This fact makes it clear that as new employees are added to the department it is profitable not only to the employees but also to the management to obtain persons who are capable of attaining higher levels of production.

Fig. 2.7 shows another distribution of the earnings of a group of operators on identical jobs (burning, twisting, and soldering) utilizing identical investments in machinery and overhead. The distribution shows that the average hourly earnings of the different operators varied from 60 cents to $1.30 per hour. The employees were working on identical machines. The differences in production (earnings) therefore *can be explained only in terms of the operators.*

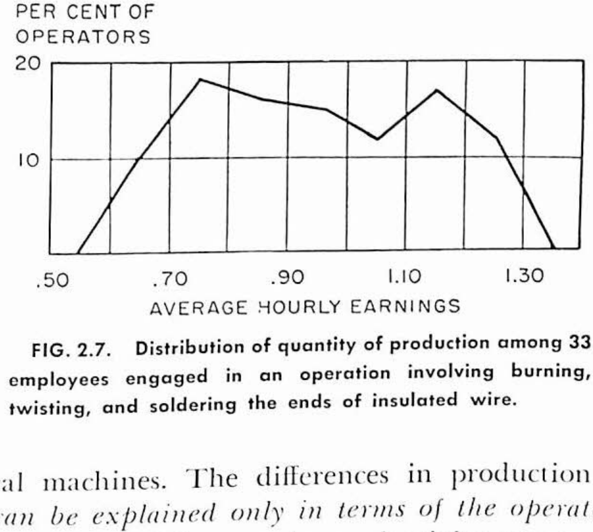


FIG. 2.7. **Distribution of quantity of production among 33 employees engaged in an operation involving burning, twisting, and soldering the ends of insulated wire.**

Still more striking differences in productivity among a group of operators are shown in Fig. 2.8. This distribution shows the production of 199 hosiery "loopers." The looping operation involves the gathering together of the loops of thread (each over a separate needle) at the bottom of a stocking after the garment is knit in order to close the opening left in the toe. Looping is a job calling for very careful and constant visual attention as well as a high degree of skill in placing the loops on the separate needles. Production on this job is ordinarily computed in terms of dozens of pairs looped per hour. Payment is on an incentive basis.

Fig. 2.8 shows a range of from .50 dozen pairs per hours (at the extreme left) to 7.00 dozen pairs per hour (at the extreme right). As might be expected, this great variation in production is due partly to differences in experience on the job. The fact that differences in experience do not account for all of the variability shown in Fig. 2.8, however, is indicated by Fig. 2.9, which shows a similar frequency distribution of production of 99 employees who had had one year or more of experience on the looping job. A careful analysis of the learning curve for looping (see Fig. 2.5 and discussion on p. 21) shows that maximum production is not reached until after a year of experience. It therefore seems reasonable to conclude that not all of the differences in production shown in Fig. 2.8 can be explained in terms of the experience factor because, when only loopers with a year or more of experience are plotted (as in Fig. 2.9), the variation in production is still from three dozen pairs to seven dozen pairs per hour.

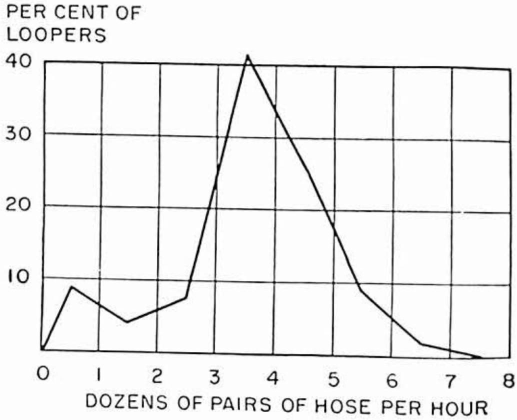


FIG. 2.8. Distribution of quantity of production in looping hosiery for 199 employees varying in experience from one month to five years.

Production data available for other jobs also show considerable

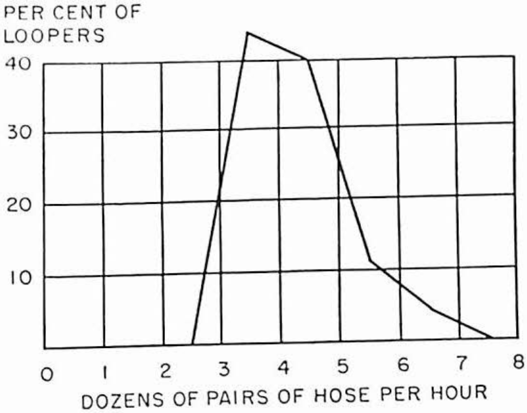


FIG. 2.9. Distribution of quantity of production in looping hosiery for 99 employees with one year or more of experience.

variability among employees. The range of performance is sometimes expressed as a ratio of the production of the least productive employee to that of the most productive employee. These ratios vary considerably, depending upon various factors, as follows: (1) the extent to which the employee has direct control over the volume of production (in the case of machine operations the individual has less influence than in the case of hand operations); (2) the extent to which the employees in the group have already been selected in terms of job qualifications; and (3) the degree of training and experience. Frequently, however, ratios of 1:2 occur, and on some jobs the ratio is as high as 1:5 or more.

Individual Differences in Merit Ratings. Merit ratings furnish another indication of individual differences in job performance. A "merit rating" is a periodic judgment of an employee by his supervisor. Usually ratings are made on each of several items, such as initiative, job knowledge, dependability, safety, and so on. One survey indicated that about 50 per cent of the companies surveyed had some type of formal merit-rating system.[15] Merit-rating systems, which will be discussed in detail in Chapter 8, usually result in a single over-all rating for each employee. This over-all rating may be taken, with certain qualifications, to indicate the employee's suitability for the job.

---

[15] "Personnel Practices in Factory and Office," *Studies in Personnel Policy*, No. 145, National Industrial Conference Board, 1954.

A distribution of the over-all ratings of 710 men from one department of a steel mill is shown in Fig. 2.10. Many cautions should be observed in interpreting the significance of these merit ratings, but it is safe to conclude that the spread of over-all ratings from 240 to 450 points suggests definite differences in the quantity and quality of service rendered by the different employees.

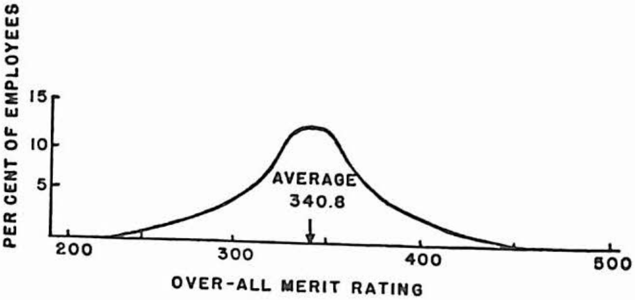


FIG. 2.10. Distribution of over-all merit ratings of 710 men from one department of a steel mill.

Individual Differences in Other Aspects of Performance. We have just seen some examples of individual differences in productivity and in merit ratings. These are only two of many different aspects of work performance. Other aspects on which people differ include quality of work, piece-rate earnings, length of time on the job (job tenure), absenteeism, and accidents. Further mention of some of these will be made later.

Stability of Job Performance. Generally speaking, the differences among employees on most aspects of performance are usually consistent differences. For example, the high-producing employee at any one time tends to remain in the high-production level, and the less efficient operator tends to remain at approximately the same low level from week to week. This fact is illustrated by the scattergram shown in Table 2.1. This table shows the relation between productivity for two successive weeks among 203 hosiery loopers. No marked shifts in productivity from one week to the other occurred. It is therefore clear that, given the production of any operator for the first week, a fairly accurate prediction of the production of that operator for the following week can be made. The coefficient of correlation[16] between the production of the first and second week computed from the scattergram in Table 2.1 was .96.

The data shown in Table 2.1 are for all employees who were on the job when the data were obtained. The experience of these employees varied from less than one month to well over two years. It will be remembered from the comparison of Figs. 2.8 and 2.9 that the variability in production of the employees with a year or more of experience is markedly less than the variability of the entire group. The correlation of .96 between production on adjacent weeks, computed from the entire group of employees, is therefore higher than would have been found if the inexperienced employees (those with less than twelve months on the job) had been excluded from the scattergram shown in Table 2.1. Even after the correction for the experience factor was made, however, the correlation between production for adjacent weeks was found to be .85.

TABLE 2.1

Scattergram Showing Consistency of Production of 203 Hosiery Loopers

The correlation between production for the two weeks shown was .96.

Average dozens of pairs per hour (second week)

Average dozens of pairs per hour (first week)	0–.4	.5–.9	1.0–1.4	1.5–1.9	2.0–2.4	2.5–2.9	3.0–3.4	3.5–3.9	4.0–4.4	4.5–4.9	5.0–5.4	5.5–5.9	6.0–6.4	6.5–6.9
6.5–6.9														1
6.0–6.4													1	
5.5–5.9											1	4		
5.0–5.4										1	11	1		
4.5–4.9								3	3	16	1			
4.0–4.4								11	16	3	1			
3.5–3.9							2	43	8	2				
3.0–3.4						1	25	3						
2.5–2.9						12	2							
2.0–2.4				1	3	1		1						
1.5–1.9				3	1									
1.0–1.4			2	1										
.5–.9		6	1											
0–.4	10	1												

---

[16] The meaning of a Coefficient of correlation is explained in Appendix A, pp. 534 ff.

These computations indicate that an employee's production level is not something that fluctuates will-nilly or that he changes daily, as he does his necktie or his shirt, but is rather a relatively fixed and permanent characteristic.

Importance of Individual Differences. Individual differences in performance, then, usually are great enough and consistent enough that they should be matters of concern both to the individuals in question and to management. Individuals whose performance is inadequate on a particular job do not have the opportunity to achieve satisfaction in their work, with possible repercussions on their total lives. Further, such performance can affect their earnings, especially if they are on piece-rate jobs. Such persons, misplaced on one job, may have potentialities for other jobs on which they could perform at more satisfactory levels.

From the point of view of management, poor work performance is an economic loss. With persons on an hourly-paid or salary basis this is easily understood, since the wages or salaries are the same per unit of time regardless of the character of work performance. What is not always as well recognized is the fact that low-performance employees also are more costly to management *even if on a piece-rate method of payment.* Although the *labor* cost per unit of production is the same regardless of how much a person produces, the *overhead* cost per unit differs. The overhead cost (capital investment in equipment, supervision, lights, other utilities, and so forth) is the same for both a high-producing employee and a low-producing employee. But if the high-producing employee produces twice as much as the low-producing employee, the overhead cost *per unit produced* is only half as much for the one as for the other.

Thus, in terms of the mutual interests of both employees and management, it should be the goal of personnel policy to place individuals on jobs in which their individual potentialities can be realized in the form of satisfactory work performance.

THE BASES OF INDIVIDUAL DIFFERENCES

Psychologists have long been interested in determining the ultimate cause of individual differences among people. Usually they have divided the major causes into the two general categories of heredity and environment. Those upholding one point of view have often minimized or even completely ignored the possible influence of the other factor. It seems probable that both factors are usually operative but that their relative importance differs markedly in the determination of different personal characteristics. Heredity would seem to be of most importance in determining such physical traits as height, weight, and strength, although it is clear that environment also has some effect upon these factors. The environmental factors of training or education are of basic importance in determining trade skill, or any other type of skill that is reached only after experience on a particular job. But to conclude that differences in skill are dependent upon differences in training does not mean that other factors have no significant effect upon the level of skill which the training produces. Whenever different levels of final performance are achieved by a group of individuals who were apparently equal when the training began, it may safely be assumed that individual differences were present at the beginning of the training, although they may have completely escaped notice. Nevertheless, they were instrumental in influencing the levels of attainment later reached by the several individuals.

An example will make this clear. It has long been standard practice in the hosiery industry to hire girls to be trained as loopers only after rigid scrutiny and interview, careful study of school records and previous employment, general physical and visual examination, and a thorough consideration of all other factors that standard employment procedure has indicated to be of importance. Obviously the girls could not be selected on the basis of their looping ability because, never having seen a looping machine, they have no skill whatever on this job at the time of employment. Except when the labor market is very tight, only those girls are hired who, in the judgment of the personnel manager, are likely with training to become good loopers. But in spite of this extensive examination and investigation at the time of employment, and in spite of the elimination of all those who for obvious reasons are clearly unlikely to become efficient on this particular job, the girls hired do not by any means become equally efficient on the operation. In fact, as Fig. 2.9 shows, after a year of experience some of the girls are more than twice as productive as others. This fact woud indicate that, though the girls might have *seemed* to be potentially equal at the time of employment, and indeed probably *were* equal in terms of all factors considered, they were not really equal at all. They differed significantly at the time of employment in one or more of the factors that determine capacity for success upon this job, but—and this is the important point—*the differentiating basic individual differences were not detected by the standard employment procedure.* It should be added that a dexterity test (described on p. 140) and certain vision tests were successful in detecting at the time of employment an appreciable number of girls who were quite satisfactory from the point of view of standard employment procedure but who did not have the basic capacities necessary for developing a high level of skill in the muscular and visual co-ordination involved in close work.

One should not infer from the foregoing illustration that the industrial psychologist recommends the elimination of standard employment procedure. He fully recognizes that a good personnel man will detect all of the obvious, and many of the subtle, reasons why some applicants should, and others should not, be placed upon a specific job. But experiments clearly show that even the best personnel men cannot determine the muscular co-ordination or dexterity of an applicant from a general interview or from an examination of his or her hands. Many if not most industrial jobs call for a combination of certain basic capacities that cannot be detected at the time of employment without special techniques and, often, standardized test procedures. These basic capacities, many of which are measurable at the time of employment, are related to future productivity, accuracy, accident proneness, promotion, versatility, tenure with the company, and many other factors that differentiate a profitable from a non profitable employee. One of the major functions of this book will be to show how these basic factors affect the development of an employee and to cite a number of illustrations of the use of selection and placement tests in industry.

CRITERIA OF JOB SUCCESS

The foregoing discussion of individual differences in work performance brings us to the consideration of what is called a *criterion* of job success. There are various circumstances in which it is useful to have some index of the relative performance of individuals. Such information is sometimes useful for administrative purposes, such as for granting wage or salary increases, or for promoting individuals to higher-level positions. In other circumstances such information is necessary for personnel research purposes. Our interest in the criterion is largely in connection with personnel research functions.

Although there are various types of personnel research in which a criterion is needed, the most common use is in the development of standards for the selection and placement of individuals on various jobs. For this purpose it is necessary to determine the kinds of characteristics that contribute to success on a given job. Although we will reserve for the next two chapters the discussion of *how* this is done, we have already seen that such a process involves the study, for one group of employees, of the relationship between personal characteristics on the one hand (frequently called *predictors*), and some measure of job performance on the other hand (the *criterion*).

Types of Criteria. It is obvious that what might serve as an appropriate criterion of job success on one job would not necessarily be an appropriate criterion for another job. Reserving for later discussion the selection of a criterion, let us now mention a few of the types of criteria that have been used in various situations. Some such criteria are *objective*, in that they are based on factual, quantitative information; others are *subjective*, in that they are based on the subjective judgments of people.

Quantity of work. For some jobs the amount of work individuals produce on their jobs is an appropriate indication of their relative job performance. The possible utility of such a criterion would, of course, depend in large part upon the availability of records of individuals' quantities of work. Typically we think of this in connection with individuals on "production" types of jobs, but the concept is also applicable to individuals on other types of jobs, such as some sales jobs and certain office jobs. Sometimes the quantity of work is reflected in secondary records, as in the case of piece-rate earnings or commissions (where these methods of payment are used).

Quality of work. This is another criterion that is frequently used, especially if records are available that reflect the quality of the work of individuals on a job. Such records may be in different forms, such as number of *acceptable units* produced, the number of *rejects* of units produced, the amount of *scrap or waste* resulting from each person's work, or the *per cent* of units worked on that are acceptable.

Learning time, or its equivalent, training cost. Other things being equal, if one new operator learns the job at a cost to the company of

$50, while another requires training that costs $150 before he reaches the same level of production and quality, the first operator is the more desirable employee.

Because of the minimum wage provisions of federal laws, it is often of great importance to economy in production to obtain trainees who are able to "earn" the minimum rate in the shortest possible time, because by law this minimum rate must be paid to the employee from the beginning of his employment.

Tenure on the job. Suppose two new employees learn a job with the same speed, and at the end of a period of time, say six months, they have reached equal levels of production and quality. If one of these employees now quits while the other stays on the job several months, or perhaps years, longer, it is clear that the employee who stays longer on the job is the better investment from the company's viewpoint. In the case of the short-term employee, it is often possible that more money is invested in training him than can possibly be realized on his low production while he is in training.

On some jobs such as that of hosiery loopers, where, according to King,[17] nine out of ten trainees fail or quit during the training period, the criterion of tenure on the job is often used as a measure of employee desirability.

Absenteeism. The percentage of working days that the employee actually reports for work has proved at times a very usable criterion. If one employee is present 97 to 98 per cent of the time and another present only 80 to 85 per cent of the time (other characteristics of the two employees being equal), the former would usually be considered the more desirable employee.

Accidents. When personnel testing programs are installed either by the safety department, the medical department, or the personnel department working in collaboration with either of the other two, one of the major objectives is to select new employees who will be relatively free from accidents. Under these circumstances, low accident experience is frequently taken as a criterion of a desirable employee.

Promotions. When a hiring program has as one of its major objectives the selection of men or women who will rapidly develop the ability to move into higher-level jobs, the speed of such upgrading may be the most satisfactory criterion to use.

---

[17] Pearl H. M. King, "Task Perception and Inter-Personnel Relations in Industrial Training, Part I," *Human Relations,* 1947, **1,** 121–130.

Job sample. When data on production and/or quality cannot be readily obtained, it is sometimes possible to obtain measurements of employee job performance by a specially set up job sample. The use of this criterion will be illustrated in Chapter 13.

Ratings. Ratings of employees by their supervisors are very frequently used as criteria. These ratings in some cases are merit ratings (which have been mentioned before); these are obtained as the result of a systematic merit-rating program. In some cases special ratings are obtained for a particular criterion purpose, such as in the development of selection and placement standards. Various rating methods will be discussed in Chapter 8. We might comment here that ratings used as a criterion may be on the basis of "over-all" job performance, or on the basis of some particular aspect of performance.

Selection of the Criterion. The selection of a criterion for any given purpose, such as in personnel research, requires careful consideration. For individuals on any given job there is, theoretically, some "ultimate" or "true" job standard by which the individuals might be evaluated. Such a standard would then provide the basis for establishing an index by which to characterize the degree of job success achieved by the various individuals.

Unfortunately, it is usually not possible to develop such an "ultimate" criterion of job success, so we typically have to select as the criterion some more readily available standard. The standard that *is* selected, however, should be as *relevant* as possible—that is, it should be relevant to the concept of what the ultimate criterion should be. Let us take a simple example. If a company is producing high-quality jewelry, presumably the ultimate criterion for certain of its jobs should embody the concept of excellence of craftsmanship. In such a case, the criterion actually selected should be one that, in large part, would differentiate among individuals in terms of the quality of their craftsmanship. Conversely, a company producing inexpensive costume jewelry would be more concerned with a quantity emphasis.

This consideration, then, suggests that the choice of a criterion for use with a particular job is essentially a management responsibility. Since management has not only the right, but the obligation, to establish the goals that are to be achieved by the organization, it is then also management's right, and obligation, to specify the standards for evaluation of employees on the various jobs in the organization.

Once one has selected the standard against which employees on a given job are to be evaluated, it is time to obtain, or to develop, the criterion values of individuals in terms of this standard. This may involve such functions as collecting and summarizing records of production, quality of work, absences, accidents, or other pertinent data, or arranging to have special ratings made.

Contamination of the Criterion. Sometimes the criterion information about individuals has been influenced by variables *other* than the abilities of the individuals. The influence of other variables is called *contamination* of the criterion. A couple of examples will illustrate the point. Let us suppose that we have two employees operating looms for weaving textiles, but that each has a different type of loom. The production records of such employees probably would *not* serve as an appropriate index of their *relative* abilities, since each individual's production is in part influenced by the productive capabilities of his own loom. Similarly, if two salesgirls in a department store are both selling ladies' handbags, but at different counters, it may be that one of the counters is in a better location than the other, thereby giving the girl at that counter a sales advantage that does not depend on her own sales ability. Although both these examples deal with two individuals, the same types of criterion contamination can occur with groups of employees.

In connection with the use of production records as criteria, Bittner and Rundquist[18] point out that "unfortunately . . . production is affected by so many things not under the control of the individual that such records rarely reflect accurately what an individual can do. They are affected by variation in the quality of the material handled, by the pace of the machine, by the pace of others in the work group, by the correlation of job assignment with length of service, and even by the correlation of job assignment with ability to do the job." An excellent example of the effect of factors beyond the control of the operator involved has been summarized by Bingham.[19] In discussing the development of a criterion for measuring the skill of bombardiers, Bingham points out that no relation was found between circular errors on alternate runs. Later analysis showed that accuracy was dependent to a very great extent upon factors other than the skill of the bombardier. Among these factors were "visibility, wind direction, bumpiness of the air, and especially the behavior of the pilot from whom the bombardier took over control of the plane thirty seconds before the bomb release." Under such circumstances, records of bombardier performances are essentially worthless as a criterion of job proficiency.

---

[18] R. H. Bittner and E. A. Rundquist, "The Rank-Comparison Rating Method," *Journal of Applied Psychology*, 1950, 34, 171–177. Reproduced by permission of the American Psychological Association.

[19] W. V. Bingham, "Today and Yesterday," *Personnel Psychology*, 1949, 2, 267–275.

Although such opinions and examples should not cause us to throw up our hands at the whole criterion problem, they should make us fully aware of the possibilities of criterion contamination.

It is not in order here to discuss the details of the methods of taking into account the influence of contaminating variables, but we might note that there are two basic approaches. One of these consists of using, in a research project, only those individuals for whom all variables *are* equal or nearly so; this approach, however, may limit seriously the number of individuals who can be used in the study. The second method consists of making some adjustment for the influence of the contaminating variable, this adjustment usually being of a statistical nature. For example, if there are ten employees working on each of two different kinds of looms, the production of each person might be related to the average production of all employees working on the same kind of loom.

Reliability of the Criterion. If a criterion is to be used as an index of the performance of individuals, it must be a *stable* index. In other words, it must reflect the relatively permanent, or continuing, level of performance of the several individuals on the job in question. This stability or consistency of performance is called the *reliability* of the criterion.

Among various ways in which the reliability of criteria can be determined, a very common method is that of comparing the criterion values for individuals for two periods of time. Such a comparison is shown in Table 2.1 (p. 29). That table shows the consistency (in other words, the reliability) of production records of 203 hosiery loopers for two time periods. It is customary to express the degree of relationship by means of a coefficient of correlation. In this instance the correlation was .96, which indicates a very high degree of stability of work performance. Generally speaking, reliability correlations in the .90's would be considered to be exceptionally good, and those in the .80's would normally be acceptable.

Combining Criteria. Sometimes there may be two or more criteria that are "relevant" to a particular job. For example, it may be appropriate to use measures of both quantity and quality of work, since a "good" employee might well be the type of individual who is high on both of these variables. If two or more criteria are available, or can be developed, and if they are all relevant, it may be well to combine them into a single criterion. The combining of two or more criteria into one usually requires some decision regarding the relative weights of the individual criteria. These judgments are based essentially on consideration of the relevancy of the individual criteria. The procedures for combining two or more criteria, with specified weights, will not be discussed here. Such procedures are included in Appendix A, pp. 528–532.

When several criteria are considered to be relevant, it is sometimes possible to combine them into a single, over-all criterion expressed in dollar units. The procedures for combining criteria in this manner have been described by Brogden and Taylor.[20]

Importance of the Criterion. Any personnel research project or personnel program that requires a criterion of job success will, in large part, succeed or fail in terms of the adequacy of the criterion used. Having a relevant, reliable, and uncontaminated criterion will not assure success of the venture, but *not* having such a criterion *will* insure its failure.[21]

---

[20] H. E. Brogden and E. K. Taylor, "The Dollar Criterion—Applying the Cost Accounting Concept to Criterion Construction," *Personnel Psychology*, 1950, 3, 133–154.

[21] Excellent discussions of criteria are available in the following publications: B. F. Nagle, "Criterion Development," *Personnel Psychology*, 1953, 6, 271–289; and J. L. Otis in Chapter V of W. H. Stead, C. L. Shartle, *et al.*, *Occupational Counseling Techniques* (New York, American Book Company, 1940).

3

THE INTERVIEW AND RELATED METHODS OF EMPLOYMENT

Employment directors recognize many of the facts concerning individual differences discussed in the preceding chapter. If an applicant has training or successful work experience in a certain area, every effort is made to place him on a job where both he and the company will profit from the utilization of his particular experience and skills. Likewise, if a job requires a person with certain physical, mental, or personality characteristics, a man possessing the required characteristics is sought. One of the major functions of the personnel manager is to recognize individual differences in people and in job requirements, and to make every reasonable effort to match men and jobs.

Various "methods" of employment are used in matching people to jobs, different methods frequently being used in combination. These methods include the interview, the application form, references, and personnel tests. This chapter will be devoted to employment methods *other* than tests; Chapters 4 to 7 will deal with the use of tests.

Before considering specific employment methods, however, it may be well to illustrate the typical procedures used in employment offices, and also to discuss the basic methods of determining the personnel specifications of jobs (sometimes also called "job requirements").

EMPLOYMENT-OFFICE PROCEDURES

The employment methods used in various companies differ somewhat. Despite these differences, however, there are certain procedures that typically are carried out, and these frequently occur in a relatively standard sequence. A common sequence, one that is recommended by Uhrbrock,[1] is shown in Fig. 3.1. Some of the steps in this sequence are not found in some companies (such as tests, or the provision for a preliminary interview); nonetheless, we have here an illustration of most of the steps through which applicants typically progress. It will be noted that interviews occur three times in this chart—following the reception of the applicant, again after employ-

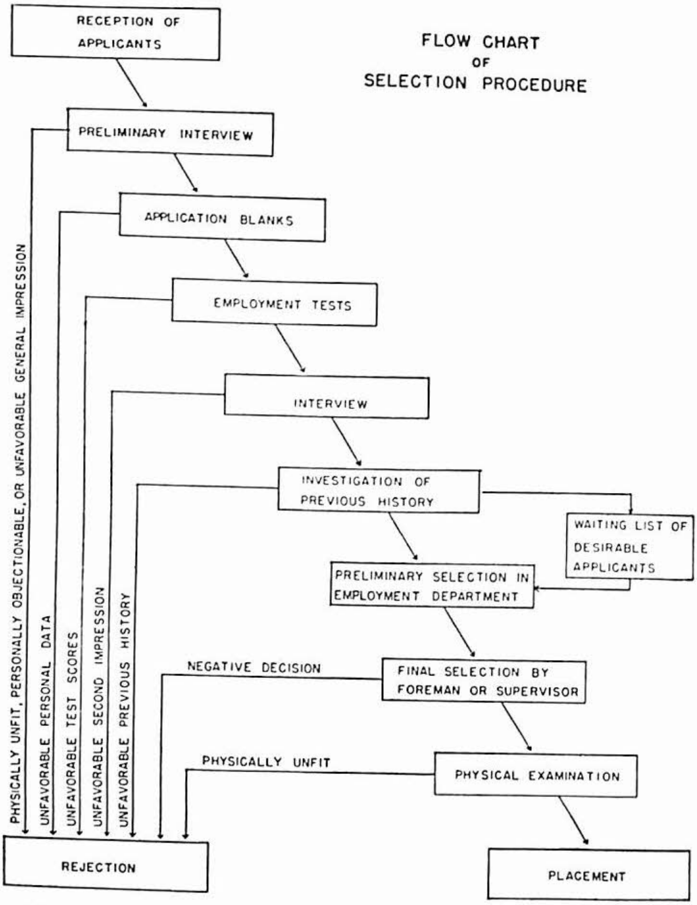

RECEPTION OF APPLICANTS
FLOW CHART OF SELECTION PROCEDURE
PRELIMINARY INTERVIEW
APPLICATION BLANKS
EMPLOYMENT TESTS
INTERVIEW
INVESTIGATION OF PREVIOUS HISTORY
WAITING LIST OF DESIRABLE APPLICANTS
PRELIMINARY SELECTION IN EMPLOYMENT DEPARTMENT
NEGATIVE DECISION
FINAL SELECTION BY FOREMAN OR SUPERVISOR
PHYSICALLY UNFIT
PHYSICAL EXAMINATION
PHYSICALLY UNFIT, PERSONALLY OBJECTIONABLE, OR UNFAVORABLE GENERAL IMPRESSION
UNFAVORABLE PERSONAL DATA
UNFAVORABLE TEST SCORES
UNFAVORABLE SECOND IMPRESSION
UNFAVORABLE PREVIOUS HISTORY
REJECTION
PLACEMENT

FIG. 3.1. Flow chart of selection procedure. (From Uhrbrock, op. cit. Courtesy the American Management Association.)

---

[1] R. S. Uhrbrock, "Mental Alertness Tests as Aids in Selecting Employees," *Personnel*, 1936, **12**, 229–237.

ment tests, and finally by the foreman or supervisor if the applicant has not been rejected before this stage is reached.

METHODS OF DETERMINING PERSONNEL SPECIFICATIONS

When someone makes a decision to place another individual on a job, he is making a prediction—specifically, he is predicting that the individual will perform effectively on the job. We all know that predictions made by human beings (including ourselves) are far from perfect. Whatever one is predicting, however—the weather, the stock market, the horses, or otherwise—his predictions become more accurate as he has more and more factual, pertinent information to use in making the predictions.

Similarly, in predicting the job performance of applicants, an employment man is able to make *better, more accurate* predictions if he is fortified with appropriate information. In particular he needs information about *what* human characteristics are related to *success* on the job in question, and information about *which* applicants *possess* such characteristics.

The characteristics or qualifications that are desired in applicants for a particular job comprise what are frequently referred to as *personnel specifications* or *job requirements*. Frequently such specifications are set up as formal standards in a job specification; they might include requirements of work experience, physical characteristics such as height and weight, minimum acceptable test scores on certain tests, and personal qualities such as appearance and verbal facility. Sometimes many such requirements exist partly or entirely in the minds of those responsible for employment. It is, of course, preferable to make such requirements as objective and as precise as possible. This raises the questions, however, of how such specifications *are* established, and how they *should* be established. Specifications are set on various bases such as judgments, job analysis, and statistical analysis.

Personnel Specifications Set by Judgments. The requirements of some jobs are determined on the basis of the judgments of people, particularly those of employment-office personnel. The requirements may, or may not, be formally specified in writing. The adequacy (validity) of the judgments varies with the job in question. If a job requires a great deal of physical effort, as in lifting or carrying heavy bags or boxes, it would be reasonable to establish some standard in terms of physical build (although one might be hard-pressed to say just *how* tall or *how* heavy a person should be for the job). Similarly, if a job requires contact with the public, such as at an information desk, it would be logical to require reasonable verbal fluency.

Many jobs, however require characteristics that are *not* as obvious as in these examples, and in such cases reliance on judgment may not be justified. How can a person "judge," for example, the kind of space visualization required in a job, or the particular kind of manipulative skill required, to say nothing about being able to judge the degree to which individual applicants possess such characteristics? Thus, there are very serious limitations to the usefulness of judgments in establishing personnel specifications.

Personnel Specifications Set by Job Analysis. Job analysis is a process of obtaining information about jobs, by such methods as observation and interview. Job analyses can serve various personnel purposes, one of them being that of setting personnel specifications. The first step in this process is to obtain a thorough job description, i.e., a detailed statement of the duties and responsibilities of the job in question. Such a job description also usually lists or describes the tools, equipment, materials, and supplies used by the worker in carrying out the job.

With information about the duties and responsibilities in mind, the job analyst then infers from these duties and responsibilities the personal characteristics that new employees should have to be placed on the job (the personnel specifications). Such inferences are relatively clear-cut for some characteristics, but are less clear-cut for others. If the job duties, for example, include the reading of micrometers and calipers, the very logical inference would be that a person should have good near vision to perform the job. Inferences about required "personality" characteristics for jobs, however, would not be as obvious.

Job specifications based on job analyses, then, are also based in large part on human judgment, although such judgments are to some degree supported by, or suggested by, factual job information.

An example of a job specification is given in Table 3.1. It will be seen that this includes both a description of the job and a listing of the personnel specifications.

Other uses of job analysis. Since we are discussing job analysis for

TABLE 3.1

Example of Job Specification

Job Title: BAKER

Alternate Title: BREAD BAKER

Industry: Hotels and Restaurants

Job Summary:

Bakes bread, rolls, muffins, and biscuits for consumption in dining rooms of hotel or restaurant, mixing ingredients to make dough, and cutting and shaping dough by hand or machine.

Work Performed:

1. Studies requirements for day and day following, and plans production in order to have bread, rolls, biscuits, and muffins freshly baked when wanted and in quantities specified.
2. Mixes (develops) dough according to recipe: Weighs and sifts flour into a bowl or mixer, adds shortening, yeast or baking powder, seasoning, and water or milk of desired temperature; either starts electric motor actuating beater that mixes and beats ingredients to form dough, or kneads dough by hand; places dough in a greased mixing bowl or proofing trough and allows it to ferment (rise or proof); may place dough in a proof box to ferment.
3. Cuts and shapes dough, sprinkling flour on work bench to prevent sticking:
    a) Flattens and distributes dough in floured pan and places it in manually operated divider which cuts the dough into sections of equal size; may cut dough to size with knife or biscuit cutter; molds all cut dough into desired shapes, by hand; may place butter, jelly, poppy seeds, or other topping on rolls.
    b) For muffins: Pours batter directly into greased muffin pans.
    c) For bread: Cuts off portion of dough with a scraper or heavy knife; weighs each portion to ensure uniformity of size; roughly shapes each portion by hand, and either places it in a greased pan or mold, or in the case of "hearth" breads (rye, French, Vienna, etc.), on a large, flat greased pan.
4. Places pans in proof box where the dough further rises. The length of time that dough is proofed depends on the recipe used (amount of yeast, water, and salt), type of flour, heat of proof box, and to some degree on the policy of the establishment or technique of the individual worker.
5. Bakes dough: Puts pans on peel and places them in oven; places hearth breads directly on hearth, using no pans; sets thermostatic temperature control or regulates fuel supply or draft to control oven temperature; may turn hand valve to inject steam into oven, moistening crust to prevent it from becoming too hard or brown; observes progress of baking, noticing the color of dough, smelling it, and pinching it or pricking it with a fork; removes baked rolls, bread, muffins, or biscuits from oven with peel or rake, or with heavy gloves.
6. Places baked products on a rack to cool.
7. Supervises a BAKER HELPER who cleans utensils and performs miscellaneous other tasks such as cutting dough.

Equipment:

Mixer; proof box; oven; proofing trough; divider; peel; rake; mixing bowls; scraper; rolling pins; biscuit cutter; work bench; refrigerator; scales; measuring cups and spoons; sifter; pans and molds for bread, biscuits, rolls, and muffins; knives.

Material:

Flour; shortening; yeast; baking powder; baking soda; milk; water; seasoning; and other ingredients used to make bread, rolls, biscuits, and muffins.

Working Conditions:

Surroundings: Inside; hot.

Hazards: Minor burns from oven or hot pans.

Specialized Qualifications:

Worker must be able to bake all kinds of bread.

Ability to bake previously prepared pastries, and make special types of breads such as cornbread and Danish pastries, may be required of worker.

Special Physical Requirements:

Strength: To lift heavy bags of flour, weighing as much as 100 pounds. Keen sense of taste and smell to determine whether ingredients are properly seasoned and sufficiently baked.

Test Standards (for use in selection of inexperienced workers for training)

Test (from USES General Aptitude Test Battery)	*Minimum Score*
Intelligence	75
Form Perception	80
Finger Dexterity	70
Clerical Perception	80

Source: Adapted from *Job Descriptions for Hotels and Restaurants*, Vol. I, U.S. Employment Service, U.S. Department of Labor, April 1938; and from "Validity Information Exchange" of *Personnel Psychology*, 1956, **9**, 383.

personnel specification purposes, it is in order to mention some of its other uses. One of these is for *training* purposes. A job analysis for training purposes usually is an extremely specific type of analysis, including a listing, in order, of the specific duties included in the job, and a description of the way in which these duties are performed. It also includes a listing of the specific knowledges and skills required, in order to provide the basis of a training program.

Job analyses are also used for *job evaluation* purposes. Job evaluation is a process of judging or evaluating jobs for use in setting wages or salaries. The type of job information used for this purpose corresponds essentially with that used for setting personnel specifications. In fact, the same job descriptions frequently serve both of these purposes.

Still another use of job analysis is for *methods improvement*. This purpose usually is served by motion and time study techniques as carried out by industrial engineers. In a broad sense these techniques can be thought of as specialized job analysis methods. On the basis of such analyses, it is sometimes possible to simplify the work layout, eliminate unnecessary motions, and make the job a simple and more convenient task to perform.

There are numerous other uses to which job analysis methods may be put, as mentioned by Zerga,[2] but the above four purposes are the most important ones.

Preparing job analyses. The process of analyzing jobs and of writing descriptions of jobs or of occupational materials requires training, experience, and attention to detail. Shartle[3] has given an excellent list of suggestions to be observed in preparing general occupational descriptions. Many of these suggestions are also appropriate for preparing job descriptions for jobs in a particular company. These suggestions are given below:

1. Always be accurate in what is expressed. The facts are what is required.

2. In describing an occupation omit words which are worker attitudes unless it is specified that you are reporting attitudes of workers. Such words include monotonous, uninteresting, interesting, pleasant, distasteful, and the like. Whether a job is monotonous or interesting depends upon the worker; it is not the occupation. One person may find being an executive uninteresting while another may thoroughly enjoy a repetitive task. The writers of occupational information are usually "white collar" workers who may tend to read their own attitudes into the job descriptions.

3. If you are describing an occupation or field of work be sure that you cover it entirely unless you indicate otherwise. That is, do not describe one phase of an occupation and give the impression to the reader that you are covering it all.

4. It is permissible to write occupational information in an interesting fashion but do not try to make the job interesting beyond giving the facts; otherwise persons may be attracted to the occupation who are not really interested in it.

---

[2] J. E. Zerga, "Job Analysis; a Resumé and Bibliography," *Journal of Applied Psychology*, 1943, 27, 249–267.

[3] C. L. Shartle, *Occupational Information* (2nd ed.; Englewood Cliffs, N.J., Prentice-Hall, Inc., 1952), p. 77.

5. Write simply and explain technical terms. Most readers will not be nearly as familiar with the job or occupation as you are.

6. Use a carefully chosen systematic outline of major and minor topics to be covered. This applies whether the material is a short description or a lengthy monograph.

7. Write concisely and give sufficient detail to meet the needs, but never pad and do not skimp on important items.

8. Try out examples of write-ups on the persons who will use the information and obtain their suggestions for improvement.

9. Follow up the use of the descriptive materials and constantly attempt to improve them.

10. Put the date of completion on each description. Revise it as often as changes in jobs and the occupation require. In using source material indicate the date it was gathered.

POSITION HANDICAP	MILLING AND BROACHING MACH. OPERATOR	LIGHT DRILL PRESS	ENGINE AND TURRET LATHE OPERATOR	BENCH LATHE OPERATOR	HEAT TREAT (INDUCTION WORK)	DISPATCHER	CRIB ATTENDANT	BURRING STRAIGHTENING	DELTA BURR MACHINE OPERATOR	BODINE MACHINE OPERATOR
BLIND BOTH EYES										
BLIND ONE EYE							X			
AMPUTATION HAND OR ARM									Lx	
DISABILITY HAND OR ARM						X				
DEAF MUTE										
DEAF										
AMPUTATION ONE OR MORE FINGERS			X					X		
AMPUTATION FOOT OR LEG	X	X					X	X		X
DISABILITY FOOT OR LEG	X			X	X	X	X	X		X
DISABILITY BOTH LEGS										
INJURED BACK OR SPINE										
CARDIAC										
LIGHT WORK RECOMMENDED										

FIG. 3.2. Job analysis for identifying jobs that can be performed by partially disabled men. (Courtesy the International Business Machines Corporation.)

Different types of occupational materials have been prepared by various private and governmental agencies, especially by the U.S. Employment Service. Many of these materials are illustrated by Shartle,[4] along with information regarding their sources.

Jobs for the disabled. One of the facts brought out by a satisfactory job description is whether the job can be performed by a man with any given physical handicap. Many plants have prepared check lists of their jobs against the more common types of physical disability, so that the employment interviewer has before him an exact statement of the physical disabilities that can safely and practicably be allowed in a man on a particular job. An example of this sort of job analysis is illustrated in Fig. 3.2. This chart and the job analyses on which it is based were prepared by the International Business Machines Corporation. The chart in its entirety includes information for 40 different jobs. In the part of the chart reproduced here for illustrative purposes, the occurrence of an **x** in the column for a particular job indicates that the specified physical handicap may be permitted in a man on that job.

The complete analysis of 40 jobs revealed that every one of these jobs would permit at least one of the specified physical handicaps, and that many of the jobs could be performed properly by employees with several of the disabilities. For example, seven physical handicaps could be allowed in an assembler, and eight in an operator on a certain part of a coil assembly.

Information of this sort is of value as a means of increasing the potential pool of satisfactory employees. To the employees and to the community it is also of value in identifying suitable work for the many persons who, although their livelihood depends upon work, in many instances have been denied employment because of arbitrary (and sometimes unreasonable) standards of physical fitness. Wagner[5] reports that although disabled persons improperly placed were poor producers and slow workers, those properly placed were equal to, or better than, nondisabled workers performing identical work. Lavos[6] cites studies showing that disabled personnel are as efficient and capable as nondisabled personnel. It is assumed that care has been taken in these situations to make the job placements in the light of the kind of job-man analysis illustrated in Fig. 3.2. A similar point of view has been adopted in the extensive work of Hanman,[7] who clearly outlines the procedures to be followed in analyzing jobs for their physical requirements and summarizes the results for a wide variety of jobs.

---

[4] *Ibid.*

[5] T. Wagner, *Selective Job Placement: a Plan for Promoting Personnel Proficiency* (New York, National Conservation Bureau, 1946).

[6] G. Lavos, "The Work Efficiency of the Disabled: An Analysis of the Available Reports on the Job Efficiency of Physically Disabled Workers in Industry," *Journal of Rehabilitation,* 1947, **13,** No. 2, 3–12, 23.

Personnel Specifications Set by Statistical Analysis. Another method of determining personnel specifications is by statistical analysis of personnel data—as for example the data that typically are available on application forms. This method involves the analysis of such data for a group of present or former employees on the job in question, the data being analyzed in relationship to some criterion of job performance.

This method is illustrated by an investigation of Tiffin, Parker, and Habersat.[8] The turnover of employees in one department of an optical manufacturing company had reached rather serious proportions. Records of two groups of employees were chosen for study. One group consisted of 27 employees who were still on the job nine months after employment. A period of nine months was chosen because any reasonably satisfactory employee remaining on the job for at least this length of time more than repaid the cost of hiring and training him on the job. A second group consisted of employees who left the job, often without notice, within three months after employment. The results of the analysis of personnel data for the two groups of employees are summarized in Table 3.2.

When comparisons of the sort shown in Table 3.2 are made by the personnel psychologist, he invariably subjects the data to a standard statistical test to determine the probability of the obtained results having occurred simply by chance. This statistical test results in a "level of confidence" that can be attached to the results. If a difference is found to be significant at the 1 per cent confidence level (as in the case of the difference in age of the "long" and "short" tenure groups in Table 3.2) we are assured that a difference as great or greater than

---

[7] Bert Hanman, "Matching the Physical Characteristics of Workers and Jobs," *Industrial Medicine*, 1945, 14, 405–430.

[8] J. Tiffin, B. T. Parker, and R. W. Habersat, "The Analysis of Personnel Data in Relation to Turnover on a Factory Job," *Journal of Applied Psychology*, 1947, 31, 615–616.

TABLE 3.2

Analysis of Personnel Data for "Long" and "Short" Tenure Employees Hired for a Specific Factory Job

Personnel Data	*"Long" Tenure Employees* (9 months or over)	*"Short" Tenure Employees* (under 3 months)	*Difference Between "Long" and "Short"*
Age	30.8 years	25.7 years	+5.1 years*
Years of Education	9.7 years	10.6 years	— .9 years†
Height	68.6 inches	69.1 inches	— .5 inches
Weight	163.4 pounds	158.3 pounds	+5.1 pounds
Marital Status	59% married	37% married	+22% married†
Number of Dependents	1.68 dependents	.72 dependents	+ .95 dependents*

* Significant at 1% level or less.
† Significant at 5% level or less.

the difference found could have occurred by chance alone only one time in a hundred. Thus the probability is 99 in 100 that there is truly a difference in average age between the long- and short-tenure employees. Noting the obtained differences found to be at the 1 per cent and 5 per cent levels of confidence, we can say that the chances are extremely great that the long-tenure employees on this job will be generally older and have more dependents, and that the chances are 95 in 100 that they will have had less formal education and that a larger percentage will be married than the short-tenure employees. Other job qualifications being equal, the employment interviewer would thus attempt to fill the job with men who are older, have dependents, are married, and have had a minimum of formal education. A later study in this plant showed that men fulfilling these personnel characteristics did remain longer on the job than men who did not fulfill them.

The obvious interpretation of facts brought out in Table 3.2 should not be applied indiscriminately to plants in all localities because labor and personnel conditions may, and frequently do, vary greatly from one locality to another. But for the time and locality in question, it seems rather clear that the employment man should seek for this job men who, when compared with the average, are older, have more dependents, have had less formal education, and are married.

Another example of this type of analysis comes from a study by

Mosel[9] of women salesclerks in a department store. In this study a criterion of sales performance was used, specifically an index of "selling cost per cent" which was derived for each salesclerk by dividing total selling cost (salary and commissions) by total net sales (dollar value of actual sales). For each of 85 departments two salesclerks were selected for the study, one being a high-criterion salesclerk (having the lowest selling-cost index) and the other being a low-criterion clerk (having the highest selling-cost index).

An analysis was then made of 42 items of information on the application forms of the 85 high-criterion and 85 low-criterion salesclerks, to determine which items, if any, differentiated between these two groups. Twelve of the items were found to differentiate. The "ideal" high-criterion (low-selling-cost) salesclerk had the following characteristics: between 35 and 54 years of age, 13 to 16 years of formal education, over 160 pounds, five years or less on next-to-last job, over five years on last job, lives in boarding house, minor executive as principal previous experience, between 59 and 62 inches in height, one to three dependents, widowed, and no lost time in two years. These items were used as a basis for deriving a "score" for each of 100 other employees, in order to double-check the results. This second group included both high-criterion and low-criterion salesclerks.

The results of this study are shown in Fig. 3.3. Specifically Fig. 3.3 shows the per cent of salesclerks with weighted scores on these 12 items *at and above* various cut-off points who were in the high-criterion group. This gives the best estimate of the per cent of high-criterion salesclerks who *would* be hired if various cut-off scores were used as a basis for employment. It can clearly be seen that persons whose scores were high (say 40 and above) would stand a much better chance of being successful than those scoring low.

From the above two examples it can be seen that the statistical analysis of personnel data can be useful with various types of criteria of job success, these two examples being concerned with criteria of tenure and of sales performance. In general, however, the analysis of personnel data has been found more frequently to be related to such criteria as tenure than to criteria more strictly of job performance.

This type of analysis frequently yields clear-cut information concerning personnel specifications that is not apparent from job analy-

---

[9] J. N. Mosel, "Prediction of Department Store Sales Performance from Personnel Data," *Journal of Applied Psychology*, 1952, **36**, 8–10.

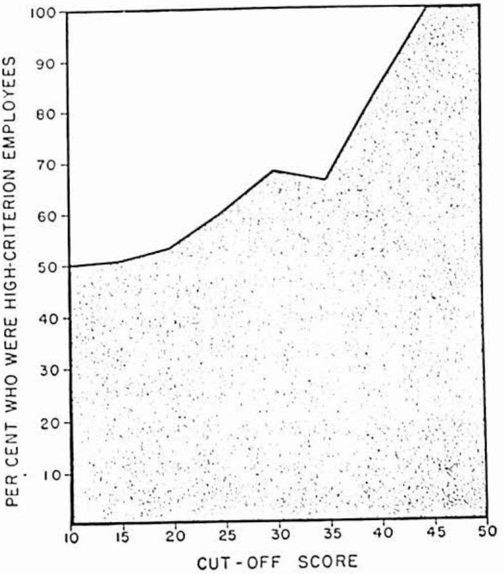


FIG. 3.3. Per cent of salesclerks at and above various cut-off scores on twelve application form items who were in high-criterion group (low selling cost indexes). (Adapted from Mosel, op. cit.)

ses or from casual observation of employees on a job. Although the examples given deal with personnel data from application forms, statistical analyses can (and should) be made of other types of information obtained from other sources, such as data obtained from the interview or from personnel tests. Depending upon the nature of the information, however, different kinds of statistical techniques may be used. The next chapter will deal with general principles and procedures in the use of personnel tests in establishing personnel specifications.

Consolidating Personnel Specifications. We have seen that the personnel specifications for jobs can be established on the basis of judgments of people, on the basis of job analysis, or on the basis of a statistical analysis of personnel data or tests. In setting specifications for any given job, however, it is not necessary to depend exclusively on any one method, but rather it is possible (and frequently desirable) to draw from these various methods. Generally speaking, it should be the policy of a personnel department to attempt to achieve relatively objective specifications that have been verified by statistical analysis. Where it is necessary to rely upon judgments in setting the specifications, one should proceed in an orderly way to make these judgments as sound as possible, preferably by the collective decision of various people who have intimate knowledge of the job in question.

It should be added that the specifications for jobs sometimes have to incorporate specific standards that may have little bearing on job performance, but rather stem from separately established policies or regulations. State and federal laws, for example, limit the employment of women and minors, and general company policy (well-founded or not) sometimes adds further employment restrictions, such as lower or upper age limits, education, and the like.

Uniqueness of Job Specifications. In the development and use of job specifications one needs to keep in mind that the standards established for the placement of individuals on a particular job *cannot* be used indiscriminately in other circumstances. The characteristics that make for success on one job may *not* be associated with success on another in the same organization, or on a corresponding job in another company or community. They will not necessarily remain constant even for the job in question over an extended period. This is due to possible differences in labor-market conditions, and even to differences in the character of what might appear to be similar jobs.

INFORMATION ABOUT THE APPLICANT

The personnel specifications of a job designate the characteristics or qualities that applicants should have if they are to be employed for the job in question. The essential function that is accomplished by the use of the various employment methods is that of obtaining information about applicants that is then evaluated against the personnel specifications, in order to be able to assess each applicant's potentialities for the job (or jobs) for which he is being considered. Assuming the validity of the personnel specifications, it is then in order to consider the extent to which the various employment methods result in information that serves the intended purpose.

The Interview. Even though many studies have been published showing that information obtained through the interview is often inaccurate,[10] that interviewers of presumably equal competence sometimes disagree markedly in their judgments of applicants, and that the interviewer's judgment of an interviewee is often unduly influenced by the latter's physical characteristics, the interview as a basic method still continues to be widely used. This continued use should strongly suggest that the interview, in spite of its limitations, occupies an important place in the majority of employment offices.

---

[10] S. A. Rice, "Contagious Bias in the Interview," *American Journal of Sociology*, 1929, 35, 420–423.

It is difficult, if not impossible, to reach a general conclusion concerning the value of the interview as an employment technique. Different studies on the value of the interview have shown markedly different results. For example, Pitney[11] reports that although only 29 per cent of men selected at random for an aircraft warning unit training center completed the course successfully, 84 per cent of those selected by interview completed the work satisfactorily. On the other hand, Dunlap and Wantman[12] found that interviews did not add to the prediction of learning to fly among aircraft pilots.

Even with great variations in the value of the interview, and in spite of the possibilities of error inherent in the interview as a technique, it has certain advantages that undoubtedly account for its long-continued use. It is sound industrial relations for every jobseeker to have an opportunity to speak at first hand with his prosspective employer (or a representative of his prospective employer) and to be given the personal attention that an application form or employment tests naturally cannot give. Because of these basic considerations, it is advisable to correct as much as possible the attendant weaknesses of the interview, rather than to eliminate it, especially since it provides an opportunity to obtain the good will of the applicant (and, through the applicant, of the community) for the company.

Probably the major deficiencies of the interview stem from one of three sources: from the method of interviewing, from the interviewer, or from the interviewee. Efforts to improve the interview as a method of employment presumably should then be centered around these three aspects.

The interview method. Many employment interviews are haphazard in that there is little planning or organization to them. Where the interviewer does not generally plan and control the direction of the interview, the subjects covered depend more on chance factors than on design. The type of information obtained for one applicant may then be somewhat different from that obtained for another. The evaluation of potentialities of various applicants cannot then be made on the same relative basis.

---

[11] R. W. Pitney, "Validity of the Placement Interview," *Personnel Journal*, 1947, **26**, 144–145.

[12] J. W. Dunlap and M. J. Wantman, "An Investigation of the Interview as a Technique for Selecting Aircraft Pilots," CAA Airman Develpm. Div. Rep. No. 33, 1944 Publ. Bd. No. 50308 (Washington, D.C., United States Department of Commerce, 1947), 59.

The deficiencies that arise from this haphazard interview approach have led to the development of what are variously called *patterned*, *structured*, or *standardized* interview procedures. Such an interview is one that follows a pre-established pattern covering certain specific areas that have been found to be related to job success. Such interviews are based on the simple premise that the best basis for judging what a person will do in the future is to know what he has done in the past. A patterned interview is organized to obtain such information in an expedient and systematic manner. The basic characteristics of an interview of this type are:

1. The interviewer works from definite job specifications.
2. He has a plan and knows what questions to ask.
3. He is trained in the techniques of conducting an interview.
4. Prior to the interview, he has checked with outside sources (previous employers, schools, and so forth) and already knows a great deal about the applicant.

Ordinarily, a printed form is used to help the interviewer adhere to the pattern. One of the most widely used patterned interview procedures has been developed by Fear and Jordan.[13] The procedure that they describe covers seven specific areas, with a number of questions under each area. These are:

1. *Previous Experience*
    a) Similar job duties?
    b) Required hand and machine tools?
    c) Same type materials?
    d) Similar working conditions?
    e) Same degree of supervision?
    f) Shown development on the job?
2. *Training*
    a) Sufficient formal school education?

---

[13] R. A. Fear and B. Jordan, *Employee Evaluation Manual for Interviewers* (New York, Psychological Corporation, 1943).

- *b*) Best-liked or least-liked subjects related to job requirements?
- *c*) Required mechanical, mathematical, or other specialized training?
- *d*) Required "on the job" training?
- *e*) Any special training since leaving regular school?

3. *Manner and Appearance*
    - *a*) Favorable, unfavorable mannerisms?
      (gestures, facial expressions, speech)
    - *b*) General appearance satisfactory?
      (features, poise, dress, personal hygiene)
    - *c*) Evidences of cultural background?
      (speech, courtesy, interests)
    - *d*) Voice and speech acceptable?
    - *e*) Physical qualifications adequate?
      (height, weight, stamina)
    - *f*) Any physical disabilities?
    - *g*) Appear nervous, high-strung?
    - *h*) Appear aggressive, self-confident?
4. *Sociability (Teamwork)*
    - *a*) Any job experience requiring special teamwork?
    - *b*) Participate in school social activities?
    - *c*) Take part in community affairs?
    - *d*) Engage in any group recreation?
    - *e*) Interests reflect liking for people?
    - *f*) Appear friendly, the kind of person who can get along with others?
5. *Emotional Stability*
    - *a*) Friction with former supervisor?
      ("chip on shoulder" or "sour grape" attitude)
    - *b*) Unsound reasons for leaving jobs?
      (incompetence, quick temper, inflexibility)
    - *c*) Unsatisfactory job stability?
      (easily dissatisfied or discouraged)
    - *d*) Reasons for leaving school?
      (reaction to failure: frank or defensive)
    - *e*) Difficult adolescent period?
      (parents divorced, all work—no play, etc.)
    - *f*) Lonely, poorly balanced life now?
      (inadequate social contacts, etc.)
6. *Maturity*
    - *a*) Work after school or summers?
      (earliest contribution to family income?)
    - *b*) Decisions dominated by family?
      (lean on family for moral support)
    - *c*) Ever lived away from home?
      (had to make own way)
    - *d*) Ever handle more than one job at a time?
      (economic drive)
    - *e*) Good sense of responsibility?

(considered transportation, wages, hours, family, in making change of jobs; how much life insurance)

f) Why did he apply for work here?
(any logical occupational goal)

7. *Leadership Capacity*
    a) Ever had leadership experience?
    (in school, former job, community)
    b) Does he want to be a leader?
    (why)
    c) Seem like natural leader type?
    (dominate or inspire confidence, respect)
    d) Reasonably aggressive, self-confident and self-sufficient?

Space is provided on the form for an evaluation of the applicant as above average, average, or below average on each of the above characteristics in comparison with employees *now on the job in question.*

The patterned interview need not be based only on the characteristics used by Fear and Jordan. Wonderlic[14] has developed a form covering four general areas: *Work History, Family History, Social History,* and *Personal History*. Several specific questions in each of these areas are provided to guide the interviewer in obtaining the relevant facts. This interview procedure, called the *Diagnostic Interviewer's Guide,* is used by the interviewer as the basis for giving ratings, which are converted to interview "scores."

In one study using this interview procedure the investigators obtained interview scores of about 300 individuals who had been hired by one company. At a later date a follow-up survey was made to find out what happened to these employees. The original group was divided into three groups—those who were still on the job, those who had resigned, and those who had been dismissed. Comparisons for these three groups were then made against their original interview scores. Fig. 3.4 rather clearly shows that more persons with high scores remained on the job than did those with low scores, and that the per cent who had to be dismissed was much lower for those with high scores.

Another patterned interview form described by McMurry[15] covers the applicant's *work record, service record, schooling, early environ-*

---

[14] C. I. Hovland and E. F. Wonderlic, "Prediction of Success from a Standardized Interview," *Journal of Applied Psychology*, 1939, **23**, 537–546.

[15] R. N. McMurry, "Validating the Patterned Interview," *Personnel*, 1947, **23**, 263–272.

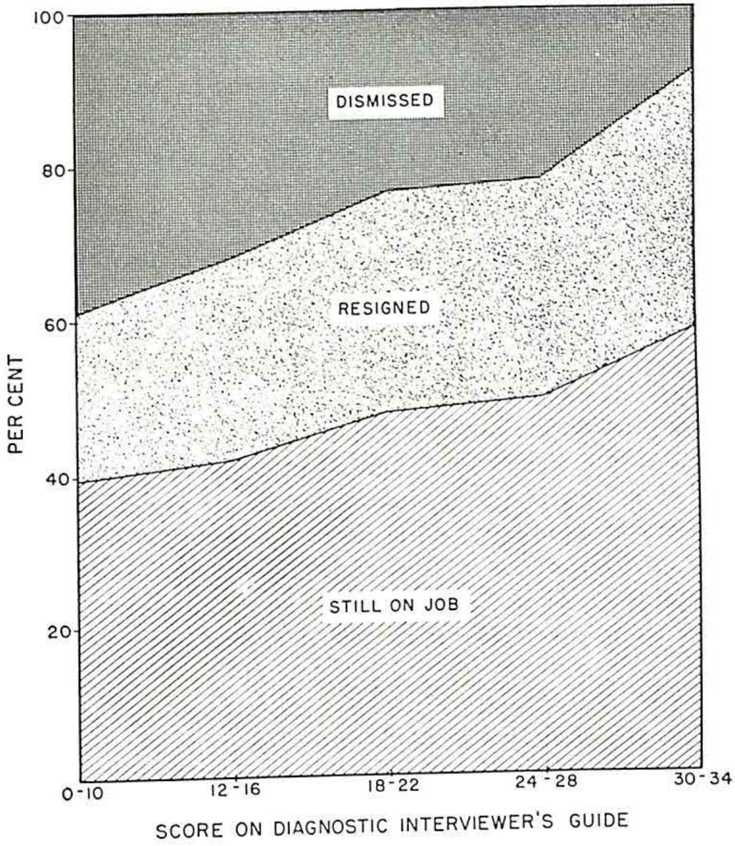


FIG. 3.4. Per cent of employees still on job, resigned, and dismissed, in relation to scores on Diagnostic Interviewer's Guide. (Adapted from Hovland and Wonderlic, op. cit.)

ment, present financial situation, domestic situation, and *health.* McMurry emphasizes that the patterned interview is essentially a fact-finding procedure, combining information obtained from the applicant with data received from schools and previous employers.

McMurry[16] summarizes results obtained with the patterned interview in several companies. A study at the Link-Belt Company in Chicago showed that interviewers' ratings of 587 men from a patterned interview were significantly related to tenure of the employees on the job. A further study of 407 men who remained on the job in this company showed a significant relation between interviewers' ratings and subsequent success on the job. Similar results were obtained in studies conducted with truck drivers employed by the White Motor Company and operators employed by the York Knitting Mills. These studies reported by McMurry show that a patterned interview, designed to accomplish a definite objective, is far superior to the random type of conversation that often occurs when undirected and/or untrained people are permitted to handle the employment interview.

---

[16] *Ibid.*

Patterned interview procedures differ in the degree to which they provide for structuring of the interview. In general, however, they do not convert the interviewer into an automatic question-asker, but rather they provide the broad framework of an interview procedure by which the interviewer is guided. It must be recognized that the patterned interview is not a panacea for all employment problems. In fact, there probably are many situations where it should not be used. But such an interview procedure has shown promise of being a useful employment device in many companies, and it is probable that its use could profitably be extended to many other employment situations.

Training the Interviewers. Interviewing is like any other skill—it does not come to one "naturally," but must be learned through instruction and practice. And like any other skill, it will not develop properly from "practice" alone if the practice is not properly directed toward the desired goal. Bingham and Moore[17] have listed a number of general suggestions that should be followed in conducting an interview, together with a discussion and explanation of each suggestion. The suggestions of Bingham and Moore that are particularly pertinent to the employment interview are summarized in Table 3.3; they can be used effectively as a guide in training employment interviewers.

TABLE 3.3[18]

Suggestions to Follow in Preparing to Conduct an Employment Interview

1. Decide just what you want to accomplish. What facts do you want to obtain?
2. Know your interviewee. Have you obtained, prior to the interview, all necessary information about him from other sources, such as the application blank?

---

[17] Adapted from W. V. Bingham and B. V. Moore, *How to Interview* (3rd ed.; New York, Harper and Brothers, 1941). 29–55.

[18] *Ibid.*

3. Make appointments. The applicant will judge the company by your reliability in making and keeping the interview appointment.
4. Provide for privacy. Remember, the interview is important to the applicant. He deserves private and courteous consideration.
5. Practice taking the interviewee's point of view. This attitude will help the applicant to overcome any tendency to be unduly nervous or apprehensive.
6. Examine and discount your own prejudices. Do not let your judgment be influenced by prejudice or bias.
7. Gain and deserve the interviewee's confidence. Unless the interviewee has confidence in you, he may not give the information you want.
8. Establish pleasant associations; they will help in establishing confidence.
9. Render your interviewee a real service. Whether or not you hire the applicant, he will appreciate sound counsel and advice.
10. Help the interviewee to feel at ease and ready to talk. Begin with a pleasant topic of mutual interest, allow time for the interviewee to become accustomed to his surroundings, and ask questions that will permit him to show pride in his knowledge.
11. Listen. The importance of this point cannot be too strongly emphasized. Remember that you learn little about the interviewee while *you* are talking.
12. Allow enough time for the interview. If you rush through it, or put pressure on the interviewee, you will not see the latter under normal conditions.
13. Do not dawdle. While the interview should not be rushed, neither should irrelevant topics be allowed to drag it out after business is over.
14. Keep control of the interview. Channel it back to the business at hand if it goes "off the course."
15. At the close of the interview, watch for additional information or new leads in the casual remarks of the interviewee. Remember that you may learn something about the applicant after you have said "good-by."

"Sampling" the interviewee's behavior. An interviewer conducting an interview obtains a "sample" of the interviewee's behavior. We all know, however, that individuals behave differently at different times, under different circumstances, and with different people. Thus, the impression that a particular interviewer receives of an applicant may not be representative of that person's typical behavior; possibly the applicant is unduly tense because of the interview, or does not "click" with the interviewer, or is not feeling well.

In order to obtain more "samples" of an applicant's behavior, on which to base a decision, it is usually good practice, where feasible, to have the applicant interviewed by two or more people, and possibly by the same interviewer on two or more occasions. In this way the people in the company who have had an opportunity to talk with the applicant can pool their impressions and evaluations with the general result that a better decision is made about the employment of the applicant in question.

Judging and rating applicants. An interviewer must make judgments about applicants—about their personal qualities, abilities, and other attributes, and their general potentialities for various types of jobs. It would be useful to see how adequate such judgments are, and also to consider methods of improving such judgments.

Adequacy of interviewer judgments. A survey of many studies of interview, or interview-like, evaluations was made by Wagner,[19] in which he summarizes information on the evaluations that have been reported for various human traits or characteristics. For each of 96 such traits or characteristics he reports the number of studies in which the trait or characteristic has been investigated, and, where reported, the *reliability* and/or the *validity* of the judgments. The reliability is the degree of consistency or of agreement between or among two or more judges; the validity is the degree to which the judgments were accurate or true judgments in terms of some separate criterion. These values are usually reported as coefficients of correlation.

Some of the results of this survey are presented in Table 3.4, which shows the reliabilities or validities of each of 13 of the 96 charac-

TABLE 3.4

Reliability and Validity of Judgments of Selected Human Traits and Characteristics in Interview-like Situations

Trait or Characteristic	*Reliability*			*Validity*		
Ability to present ideas	.42					
Alertness	.36					
Appearance	.34			.13		
Background, family and socio-economic				.20		
Energy	.64					
Initiative	.57					
Intelligence or mental ability	.96	.87	.77	.58	.82	.45
	.62	.90		.94	.51	.70
Personality				.21		
Self-confidence	.77					
Sociability	.87	.72		.37		
Social adjustment				.22		
Tact	.26					
Over-all ability	.71	.48	.24	.27	.21	.16
	−.20	.26	.43	.87	.23	
	.68	.61	.85			
	.55					

Source: Wagner, *op. cit.*

---

[19] R. Wagner, "The Employment Interview: A Critical Summary," *Personnel Psychology*, 1949, 2, 17–46.

teristics investigated. Since this information comes from a wide variety of situations, it must be accepted with reservations, but nonetheless we can see that there are marked differences in the reliabilities and validities of the various traits and characteristics. For example, the reliability of ratings on "Alertness" was .36, whereas for "Sociability" the reliabilities (in two situations) were .87 and .72. The validities ranged from .13 for "Appearance," to those in the 80's and 90's for "Intelligence or mental ability." We can also see, however, that for a given characteristic, such as "Over-all ability," the range of validities and reliabilities varied a great deal; the validity coefficients for this characteristic ranged from .16 to .87, and the reliability coefficients ranged from −.20 to .85.

Such studies, as well as experience with the interview, suggest that interviewers can make better judgments about some characteristics or traits than about others. Thus, an interviewer might be able to observe quite readily that an interviewee has a pleasing smile, effective speech and voice, and general air of confidence. It is not as easy for him to determine whether a man can operate a milling machine, read a micrometer, or find the trouble in a defective machine; or whether a girl can type accurately, add a column of figures correctly, or avoid mistakes in spelling. This fact—that the interview is a "natural" for determining certain characteristics but a rather ineffective method of determining others—should be kept in mind when one discusses the "validity of the interview." It would seem to be a mistake to make a single over-all evaluation of the interview as a method. Its value depends in part upon the job in question. For example, the very same personal characteristics that enable an applicant to make an excellent impression on the interviewer will frequently enable him to make an excellent impression on a prospective customer, *if he is hired as a salesman,* owing to the fact that the interview often taps the same personal characteristics that successful job performance as a salesman requires. But the traits that make a good impression in an interview are often far removed from those required for successful job performance as a laboratory technician, an assembler, or a machine operator. Thus a certain type of interview might be highly satisfactory in selecting employees for one job and far from satisfactory in filling another.

The use of interviewers' ratings. Although interviewers' judgments about traits or characteristics of applicants are far less accurate than would be desired, we should not conclude that they have no place in employment procedures. Rather, efforts should be made to improve such judgments, and to use them properly. One method of improving interviewers' judgments is by the use of appropriate rating scales. The ratings resulting from the Diagnostic Interviewer's Guide, mentioned earlier (p. 56), are examples. Such ratings are converted to scores, as shown in Fig. 3.4.

Another example of a rating scale used by interviewers is shown in Fig. 3.5. This form was developed to aid inexperienced interviewers in limiting their attention to those things about an applicant that can be learned during an interview. It will be noticed that each of the six characteristics listed can be well observed or evaluated during an interview.

Evaluating work history. Information concerning an applicant's work history is often of value to the employment interviewer largely because, as mentioned in the previous discussion, what a man has done in the past is a rather good indication of what he will do in the future. Many of the most accurate methods of predicting human behavior are based on this principle.

Work-history information concerning an applicant is often obtained during an employment interview. The question naturally arises concerning the accuracy and validity of information so obtained. A study by Keating, Paterson, and Stone[20] has shown that there is remarkably close agreement between work-history information obtained during an interview and the information obtained later from the stated previous employers. This indicates that, in general, the information on work histories that is obtained during the interview is relatively accurate.

The next question we would like to ask relates to the usefulness of work-history information in the *prediction* of future success of applicants. In this connection, one can consider the usefulness of *various kinds* of work experience as predictors of success on a particular type of job, and also the usefulness of previous work experience on the *same* kind of job for which a person is being considered. Where circumstances permit, statistical analyses should be made to provide information of these types.

A study of the usefulness of various kinds of work experience as

---

[20] E. Keating, D. G. Paterson, and H. C. Stone, "Validity of Work Histories Obtained by Interview," *Journal of Applied Psychology*, 1950, 34, 6–11.

NAME ______________________________ DATE ______________

How does his appearance impress you?

☐	☐	☐	☐	☐
Makes a very poor impression; slovenly, unkempt, or flashy.	Somewhat careless about clothes and appearance.	Clothes and appearance are ordinary; neither shabby nor exceptional.	Creates a better than average appearance.	Creates an excellent appearance; clothes are neat and appropriate.

How well does he talk? Does he express himself clearly and adequately?

☐	☐	☐	☐	☐
Very good expression and speech; forceful and effective.	Good expression; talks deliberately and fairly fluently.	Average expression; talks fairly well but not with great fluency.	Poor expression; tries to express himself but does not succeed very well.	Very poor expression; talks little and doesn't express self well.

How sociable and friendly is he? Does he seem to be a good mixer?

☐	☐	☐	☐	☐
Unfriendly, unsociable, or bashful.	Somewhat reserved or retiring.	Friendly but not overly expressive.	Friendly and quite expressive.	Extremely social; treats new acquaintances as if they were old friends.

How good is his work record?

☐	☐	☐	☐	☐
Excellent; never changed jobs unless there was a clear and good reason for the move.	Good; usually did not change jobs without good reason.	Average. Some change without reason and some with good reason.	Poor; several moves without apparent reason.	Very poor; frequent shifting of jobs.

How much initiative is indicated by his past jobs and work experience? Has he been a "self-starter"?

☐	☐	☐	☐	☐
Practically no initiative; has done little or nothing "on his own."	Lacking in initiative but once or twice has shown some.	About as apt to go ahead as not; performance variable.	Has considerable initiative but only when certain interests are involved.	Has a great deal of initiative; has done many things that show real desire to get ahead.

How anxious does he seem to be to work for this company?

☐	☐	☐	☐	☐
Extremely interested in the company; is completely "sold" on the organization.	Expresses an interest in the company and probably prefers it to some other companies.	Not excited about the company but might develop a feeling for the organization.	Somewhat indifferent; interested mostly in pay.	Very indifferent; just another job; might leave without notice.

FIG. 3.5. Part of a simple rating scale for use by interviewers. (From Interviewers' Aid No. 1, Interviewers' Rating Scale, Science Research Associates, Chicago, Ill.)

predictors of success on a given type of job has been reported by Steward.[21] In particular, this study dealt with the relationship of various types of prior experience to success in life insurance selling. The results are summarized, in part, in Table 3.5. For each of various prior occupations, there is shown the correlation with success in life insurance selling.

TABLE 3.5

Relation of Type of Prior Experience to Success in Life Insurance Selling

Previous Occupation	*r*	*Degree of Relationship*
Real est. salesman	+.71	
Sales manager	+.67	High overlapping
Advertising mgr.	+.35	(Favorable occupations.
Lawyer	+.22	Man qualify for in-
Social sci. teacher	+.19	surance)
Banker	+.18	
Y.M.C.A. secretary	+.18	
Pub. utility sales	+.15	Med. overlapping
City school supt.	+.13	
Personnel mgr.	+.10	
President mfg. co.	+.09	
Office worker	+.06	
C. P. A.	+.04	Low overlapping
Purchasing agent	+.02	
Author-journalist	+.01	
Y.M.C.A. phys. dir.	+.01	
Minister	+.01	
Accountant	.00	No relationship
Musician	−.05	
Artist	−.07	
Osteopath	−.10	Slight aversion
Policeman	−.11	
Production mgr.	−.14	
Doctor	−.26	Medium aversion
Architect	−.31	
Psychologist	−.32	
Dentist	−.34	
Forest service	−.36	
Math.-sci. teacher	−.41	
Printer	−.45	
Farmer	−.50	(Unfavorable occupations.
Carpenter	−.53	Few qualify for life insur-
Mathematician	−.58	ance)
Eng. (civil, mech.)	−.61	
Physicist	−.69	
Chemist	−.71	Great aversion

---

[21] Verne Steward, "Selection of Sales Personnel Reports," Bulletin 1, January 1947, 2 (distributed by Verne Steward & Associates, South Gate, California).

The results in Table 3.5 show that in selecting men for this type of selling, men with prior work experience in such fields as real estate selling, sales management, and other occupations appearing near the top of the list, have a much greater chance of success on the job than do men with previous experience in fields appearing near the bottom of the list. Previous experience in fields indicated near the center of the list is not related either positively or negatively to success in selling life insurance. Information of the sort summarized in Table 3.5 is of considerable value to an interviewer in evaluating an applicant's work history in relation to possible success on the new job.

The interviewer may be interested in predicting the performance on a given job of applicants who have had previous experience on the same type of job, or one somewhat related to it. The problem here is one of evaluating, on the basis of a record of work experience, the degree of *proficiency* of *experienced* applicants.

A study by Mosel[22] sheds some light on this. This study dealt with Civil Service employees on a variety of jobs. Prior to employment, these individuals had been rated on the basis of their work experience that was pertinent to the jobs for which they were being considered. The ratings involved the use of weights assigned according to type, level, and length of pertinent experience. These experience ratings, made prior to employment, were later correlated with criteria of actual job performance; the criteria were ratings by supervisors.

The correlations with these criteria for the several jobs are given in Table 3.6. (Later reference will be made to the correlations with the "recommendation" scores.) It will be seen that, in general, the correlations of experience ratings with the later job performance criteria are slight. The only exceptions are for auto mechanic and painter, but even though these correlations are statistically significant, they are still low. Such results tend to support the contention that although work-history information may indicate *what* an applicant has done in the past, it does not necessarily reveal *how well* he did it. In any event, one should not assume that work-history information provides a basis for evaluating the *quality* of applicants' previous job performance unless there is some particular evidence to support this assumption.

---

[22] J. N. Mosel, "The Validity of Rational Ratings on Experience and Training," *Personnel Psychology*, 1952, **5**, 1–10.

The Application Form. The application form, as the name implies, is intended to obtain information about the applicant that will reveal certain facts concerning his employability. The employability of an applicant may, and frequently does, vary from one company to another, from one community to another, and even from time to time in the same company and community. Company policy may restrict employment according to certain principles that vary legitimately from one company to another. The application form contains information that tells the interviewer whether the applicant meets the specifications stipulated by company employment policies.

TABLE 3.6

Correlations of Experience Ratings with Job Performance Criteria and with "Recommendations"

		Correlations	
Job	*N*	with Performance Criteria	with "Recommendations"
Carpenter	51	.19	.24
Equipment Repairman	40	.16	.09
Machinist (I)	76	.00	−.30 †
Machinist (II)	100	.04	.41 †
Machine Operator	108	.14	.15
Ordnanceman–Torpedo	125	.09	.02
Radio Mechanic	116	−.17	
Aviation Metalsmith	98	.12	
Highlift Fork Operator	116	−.04	
Auto Mechanic	99	.27 †	
Painter	75	.26 *	
Ordnanceman	100	−.09	−.22 *
Hand Compositor	116	.16	.27 *
Economist	41		.22
Budget Examiner	21		.63 †
Training Officer	47		.39 †

* Significant at 5% level. † Significant at 1% level.

Source: Mosel, *op. cit.*

The application form usually asks for certain facts concerning education and other formal training, previous employers, and work history, and for related information concerning the applicant's previous connections and activities.

The basic assumption underlying the use of an application form as one of the tools of the employing official is that the information obtained is significant in indicating something about the potential value of the applicant as an employee of the company. It would seem logical, therefore, that all questions to appear on the application form be carefully selected, so that only questions eliciting significant information are included. A cursory examination of any random sample of application forms in current use suggests strongly that often no such care has been given to the selection of questions included on these forms. Many forms are lengthy documents, and most employment men will admit that little or no use is made of the answers to many of the questions that appear on them. Often, in such instances, the application form consists of a compilation of questions asked on other forms, with the addition of any other questions that may occur to the members of a managerial and industrial relations committee appointed to recommend an application form.

The use of an unwieldy form consisting of a long list of questions (many of which do not yield information really related to the potential job success of the applicant) is open to criticism for several reasons. Applicants are not favorably impressed when they are asked immediately for a considerable amount of personal information about themselves that they feel (often correctly) should be of little concern to the company, inasmuch as it is not related to their qualifications for the job. A second disadvantage of a long application form, which asks for much unimportant information along with a smaller amount of significant information, is the fact that the interviewer has no means of determining what information is important. He may, therefore, arrive at a decision concerning the applicant that is unduly influenced by the latter's answers to questions that are not important and that should not have been included on the form.

The considerations discussed strongly suggest the advisability of using an application form that asks only for information really pertinent to the employability of the applicant. A satisfactory decision concerning what questions to include cannot be reached solely by examining the forms used by other companies nor by accepting blindly the judgments and suggestions of the members of a committee. A satisfactory decision on the content of an application form requires a simple but accurate statistical comparison of the later success or failure on the job of employees who give different answers to various questions.

This process requires time, but it is by far the best means of developing an application form that obtains information of real significance to the later job success of the applicant. Two examples of this approach were given on pages 49 to 50 in our discussion of the use of statistical analysis in developing personnel specifications. Further examples of its use will illustrate the application of the principle. In the development of an application form for prospective life insurance underwriters, a large number of questions were included on the form during its developmental stage. As men were hired and placed on the job (without reference to the information obtained by the form), the facts concerning which questions asked on the form were pertinent and which were not were revealed. It was found, among other things, that 66 per cent of the married men with dependents succeeded on the job, whereas only 40 per cent of the single men without dependents achieved the same degree of job success.[23] Thus, other things being equal, a married man with children has 66 chances in 100 of succeeding in this type of work, whereas a single man has only 40 chances in 100 of attaining success.

Many items of information that seemed to be of importance did not, in the trial, distinguish the successful employees from the unsuccessful. The information items finally retained on the form were therefore reduced to *those that were proved to be related to job success.* In the case of the information concerning marital status referred to in the preceding paragraph, applicants were asked whether they were married and had children *not* because someone thought this information was related to job success, *not* because the form could not be considered complete without it, and *not* because it was obtained from one or more forms that other companies had used or were using. This information was included on the form for one and only one reason—namely, that the *sales records showed conclusively that married men with children sold more insurance than single men.*

The Aptitude Index is a form developed in this manner. It calls for biographical items concerning the applicant (such as age, marital status, number of children, and amount of life insurance owned); it also includes a number of questions that deal with the applicant's interests. Every item appearing on the form was used because extensive research had shown that the answers to it were related to tenure and success in the field of selling life insurance.

---

[23] Verne Steward, *The Use and Value of Special Tests in the Selection of Life Underwriters* (Kellaway-Ide-Jones Company, Los Angeles, California, 1934), 65.

Results obtained with the Aptitude Index when used with new salesmen are shown in Table 3.7.

TABLE 3.7

Aptitude Index Score and First-Year Performance

AI Score	*No. Contracted*	*% Survive*	*Per cent "Success"**
A	570	52%	30%
B	688	44	22
C	256	42	16
D or E	28	29	11

* "Success"—one-year survival with production equal to or greater than the median production of one-year survivals in the agent's company.

Source: Based on "A Selection Study," Research Report 1951–2, File No. 424, Life Insurance Agency Management Association.

Table 3.7 clearly shows, among other things, that 30 per cent of the 570 new salesmen who made a score of A on the Aptitude Index were "successes," i.e. they survived a year and produced equal to or above the median production of all one-year survivals in their respective companies. In contrast, only 16 per cent of the men making a score of C were equally successful.

Results obtained with the Aptitude Index may also be shown graphically as in Fig. 3.6. Fig. 3.6 shows the first-year production per 1,000 new agents among agents scoring at different points on the Aptitude Index.

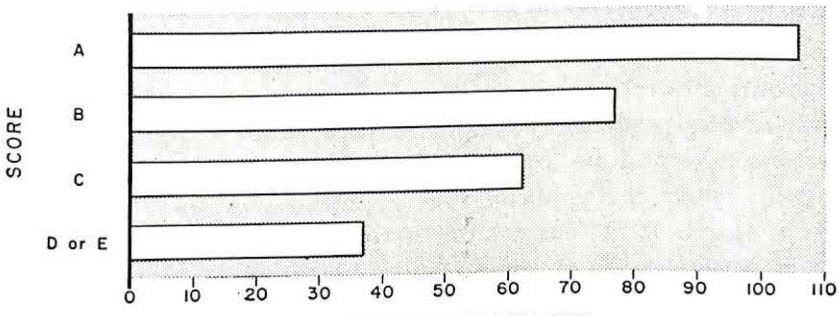


FIG. 3.6. Total first-year production from 1,000 insurance-salesman recruits at each score level on the Aptitude Index. (Courtesy the Life Insurance Agency Management Association.)

The above discussion is not intended to imply that the Aptitude Index would work equally well, or even at all, in selecting some other type of salesman.

The examples given above, plus those previously used in the discussion of statistical analysis of personnel data, have been presented in some detail to emphasize the value of a method that frequently can result in a valid application form, one that is concerned only with pertinent information.

Many companies, of course, use a single application form with applicants for a wide variety of jobs. In such cases it is not feasible to have a separately designed form for applicants for each job: rather, a single, more general form needs to be used. Even in such instances, however, the same principle applies: the interviewer should consider primarily those items of application-form information which have been found to be related to job success on the *specific* job in question.

One method that has been developed for focusing the attention of the interviewer on only such *pertinent* items of information is that of a "template" for each job which can be placed over an applicant's application form. The template for a given job has openings which reveal only those application-form items which the interviewer should consider in selecting individuals for the job in question. Fig. 3.7 shows an example of a template of this sort, together with an employment application form. When the form is viewed through the template only the items of information pertinent to the job in question are visible to the employment interviewer.

References. In employing men or women for many kinds of jobs, particularly higher-level jobs, references are sometimes used. Usually these are in the form of letters of recommendation, although they sometimes also take the form of questionnaires to be filled out, or even personal contact between the prospective employer (or his representative) and the "reference" individual (usually someone in the company where the applicant was previously employed). Theoretically it would seem that a letter about an applicant written by a former employer, or by some responsible person who has been in close contact with the applicant, would provide a highly satisfactory method of obtaining information that is accurate and comprehensive. To accomplish this result, however, the writer of the letter must be both able and willing to put into writing his full and frank

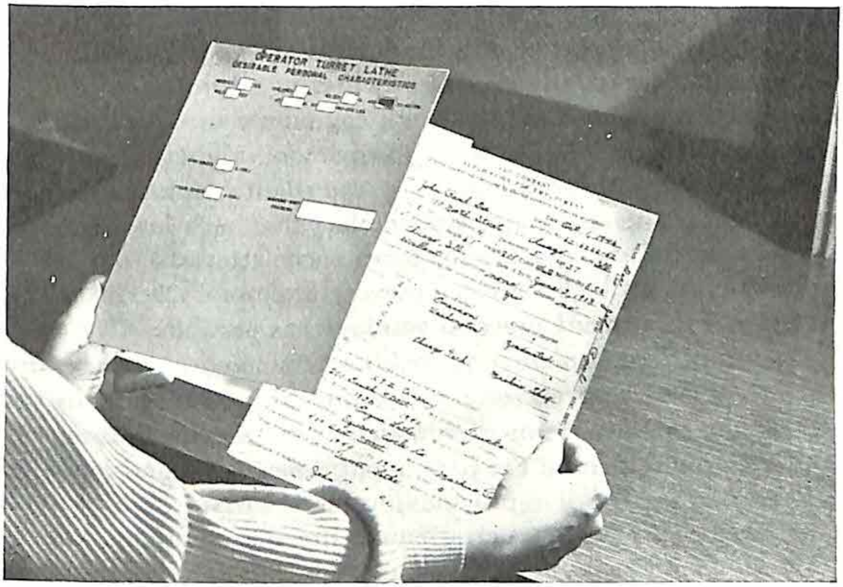


FIG. 3.7. Template for use by employment interviewer in evaluating application form for turret lathe operator. (From W. F. Wood, "A New Method of Reading the Employment Questionnaire," Journal of Applied Psychology, 1947, 31, 9–17. Courtesy the American Psychological Association.)

opinion. Such ability and willingness are not always present, and there are some circumstances under which writers of letters about applicants are inclined to give something less than a comprehensive and accurate picture of an applicant. This fact is indicated by the results of a questionnaire survey of 100 sales managers, partially summarized in Table 3.8.

TABLE 3.8

Responses of 100 Sales Managers to Questions Concerning Their Practice in Writing Letters of Recommendation[24]

Questions Asked the Sales Managers	*% Yes*	*% No*
In replying to a letter of inquiry from a prospective employer regarding the fitness of one of your former employees,		
Do you tell only the good things you can say about the man?	77	23
Do you always give the employee the benefits of any doubts?	85	15
Do you point out the man's failings and weaknesses as well as his strong points?	39	61

---

[24] By a New York Sales Manager, "To Whom It May Concern," *Sales Management*, 1923, 9.

The unwillingness of many to write a full and frank statement is probably due in part to a natural reticence to put into writing comments that necessarily become a matter of record.

Letters of recommendation are not commonly used in plant and shop employment. The hazards of this tool, therefore, need be guarded against only occasionally by the plant employment interviewer, but he should be aware of the fact that such letters are not always complete and exact, and, when such letters are offered for his consideration, he should make every attempt to verify the information from as many previous employers as possible.

The use of questionnaires probably offers some advantage over the conventional letter of recommendation, since the company considering the applicant can then solicit more specific information about the applicant from the respondent. Even such questionnaires, however, frequently fail to provide information that would be useful in predicting future job performance. In the study by Mosel,[25] previously mentioned, for example, recommendation questionnaires were available for individuals who were employed on several of the jobs covered by the study. These questionnaires had been solicited from presumably qualified references prior to employment of the individuals. They required the respondent to rate the applicant on occupational ability and character, to give his judgment on whether the applicant was especially qualified for the specialty for which he was applying, and to indicate whether he (the respondent) would hire the applicant. These items were objectively scored to yield a total "recommendation" score. These scores were later correlated with criteria of job performance after the individuals had been on their respective jobs for some time.

These correlations are given in Table 3.6 on p. 66. This is the table that also shows the correlations between experience ratings and the job performance criteria. These correlations are quite variable, some of them even being negative. It is interesting to note that the correlations for the two groups of Machinists were —.30 and +.41, both statistically significant but in opposite directions. The correlations for certain jobs—namely Hand Compositor, Budget Examiner, and Training Officer—indicate that the recommendation questionnaires might be of some utility in selecting applicants for these jobs. The general implication of the results, however, is that such questionnaires *do not necessarily* provide information that is useful in predicting future job success of applicants. In other words, before placing a great deal of reliance on such references it would be desirable to study the utility of such information for the employment of applicants for the job or jobs in question.

---

[25] Mosel, *op. cit.*

It might be added that generally greater reliance can be placed on references when the person who is serving as a reference is known to the representative of the employing concern. The further point might be added that there seems to be evidence that intensive "investigations" of applicants may provide useful information for employment purposes. In the study reported by Mosel,[26] for example, skilled investigators interrogated past employers, co-workers, supervisors, and acquaintances of applicants for the three professional jobs of Economist, Training Officer, and Budget Examiner. These inquiries were carried out to obtain facts about the applicant's experience, past job performance, character, ability, and personality. The investigators' reports were later used as a basis for ranking the applicants in terms of their qualifications for employment. The rankings were then correlated with the later criteria of job performance, with resulting correlations as follows: Economist, .33; Training Officer, .63; and Budget Examiner, .79. All of these are statistically significant at the 5 per cent level. The last two, in particular, are of such value that they would be of considerable usefulness in employment of applicants for such jobs. It should be kept in mind, however, that part of the information obtained by this type of investigation also can be obtained by interview, and used as a basis for rating applicants on experience.

Personnel Tests. Personnel tests, when given and interpreted properly, constitute one of the most effective aids available to the employment interviewer. Such tests very often quickly and accurately reveal information about an applicant's skills, abilities, aptitudes, and interests that is very difficult to obtain by any other means. This does not mean that personnel tests constitute the final answer to the employment problem. Certain characteristics of an applicant that are important in determining potential success or failure on the job are not entirely revealed by any personnel tests that have been developed up to this time. But as a supplement to other employment techniques, and as an aid to the man conducting the employment interview, personnel tests have been proved in scores of practical situations to be worth many times their cost.

---

[26] Mosel, *op. cit.*

The application of psychological methods in the development and standardization of personnel tests constitutes one of the most significant contributions of industrial psychology. The importance of this field, the great amount of attention that psychologists have given to it, the extensive research on tests that has been published in psychological and personnel journals, and—perhaps, above all —the fact that a proper use and interpretation of personnel tests requires careful study, justify the rather extensive coverage of this subject contained in Chapters 4 through 7.

4

GENERAL PRINCIPLES OF PERSONNEL TESTING

Personnel tests as utilized by modern business and industrial organizations are used for various purposes, including the following:

1. Selection of new employees.
2. Placement of employees on jobs.
3. Transfer of employees.
4. Promotion of employees.
5. Evaluating effectiveness of training programs.
6. Determining needs for training.
7. Employee counseling.

Tests are used in such organizations probably more for selection and placement purposes than for the other purposes. Although the primary emphasis in this text will be on the use of tests for selection and placement, it should not be assumed that these are the only purposes of tests. It might be added, incidentally, that much of what will be said about tests used for selection and placement purposes also would be applicable to tests used for other purposes, such as in making transfers or promotions, or in connection with training activities.

The nature and significance of individual differences among industrial employees were indicated in Chapter 2. Since every personnel man recognizes the existence of such differences, it is quite unnecessary for the industrial psychologist to re-emphasize the importance of proper employee placement. The personnel procedures in use in the modern employment office have been installed primarily for the purpose of making a careful and adequate selection and placement of employees. The psychologist's function, rather, is to bring to the attention of the personnel manager those aspects of the theory and application of modern psychology that have been found helpful in the performance of this difficult task.

GENERAL CONCEPTS

Testing Supplemental to Other Employment Procedures. Many industrialists are skeptical about the use of psychological tests in employment because they do not wish to risk a change from present methods that are known to be reasonably satisfactory to new methods that have not been thoroughly tried in actual practice. Their skepticism is clearly justified. As Cleeton[1] has emphasized, personnel tests should always be considered a supplement to, not a substitute for, other employment procedures. No psychologist who has thought at all about the problems of modern industry would seriously suggest that present employment procedures should be eliminated. Indeed, many of the advances that modern industry has made during the past forty years can be traced directly or indirectly to the procedures now in use in every modern industrial employment office. But the fact that these methods, though excellent in many respects, are still not perfect is proved by the marked individual differences among employees that any study of differential production will reveal. A considerable amount of research, both in industry and in the laboratory, has shown that still further improvements in employment methods can be attained when psychological tests and methods are used as supplements to other employment procedures. Let us keep in mind this word "supplement."

Extent of Use of Employment Tests. Probably the most comprehensive survey on the use of employment tests is one conducted by the National Industrial Conference Board.[2] The results of this survey of about 3,500 companies are reported in Table 4.1. It can be seen that the use of tests is distinctly related to size of company. About half of the large companies (5,000 employees or more) used employment tests.

---

[1] G. U. Cleeton, "Use of Tests in the Transit Industry," *Passenger Transportation*, 1946, 4, No. 19.

[2] "Personnel Practices in Factory and Office," *Studies in Personnel Policy*, No. 145, New York, National Industrial Conference Board, 1954.

Selection or Placement as the Function of Tests. Employee tests have often been considered primarily as devices to aid in the selection of employees. This viewpoint is followed by the assumptions that (1) when tests are used systematically, many applicants will be rejected—that is, not employed at all—and that (2) unless there are significantly more applicants than there are jobs to be filled, the testing program loses its effectiveness.

It is unfortunate that the selective rather than the placement features of employee tests have been given the greater emphasis in most discussions of this subject. Perhaps this is because many industrial testing programs were inaugurated during the depression years of 1930 to 1935, at which time there was an unusual abundance of applicants for nearly every job. But in a period of prosperity there is usually not an abundance of applicants. Indeed, many industries find it necessary to relax all employment standards, including not only psychological test requirements (in industries where these have been established) but also standards dealing with medical and physical requirements, age, and marital status. When a testing program has been installed on the basic premise that it is a *selection* program, the fact that it is also an excellent *placement* program is often forgotten. Yet the value of a testing program in placing employees during a period of rapid plant expansion is fully as great as its value in selecting employees during more static periods. Selover and Vogel[3] aptly summarize the matter by stating that "a testing program can be used to advantage even when the shortage of labor requires that every applicant for employment be hired to fill some position. Under such circumstances, provided there is more than one

TABLE 4.1

Per Cent of Companies Reporting the Use of Employment Tests, by Size of Company

Size of Company by No. of Employees	*Per Cent Using Tests*
5,000 and over	53
1,000 – 4,999	43
250 – 999	28
Under 250	15

Source: Adapted from "Personnel Practices in Factory and Office," *Studies in Personnel Policy* No. 145, New York, National Industrial Conference Board, 1954.

---

[3] R. B. Selover and J. Vogel, "The Value of a Testing Program in a Tight Labor Market," *Personnel Psychology*, 1948, **4**, 447–456.

kind of job open, the testing program enables us to place the applicant in that job for which he appears to be most suited."

One of the most effective uses of a testing program that the writers know of was in a war industrial plant that was hiring, as the medical director stated, "anyone healthy enough to walk into the plant." In this plant the sole purpose of the tests was *placement* of the many new employees on the specific jobs to which each was best adapted. The value of tests is by no means limited to situations in which there is an abundance of job applicants.

Psychological Tests Are Not Infallible. A further point that the advocate of testing procedures should make clear is his recognition of the fact that psychological tests are not infallible; that they sometimes give results that are not a true indication of the potential job success of the applicant. Any new procedure, whether in employment, production, advertising, or whatever, should be evaluated not in terms of whether it achieves perfection, but whether it results in some useful degree of improvement over methods that have preceded it. Thus, if the labor turnover in a given department has been 25 per cent per year among employees placed by previous methods; if it is found that new employees placed by psychological tests consistently show a turnover of only 20 per cent per year (all other factors indicating a desirable employee remaining constant); and if the expense of administering the testing program is less than the amount saved by the reduction in labor turnover, the testing program would ordinarily be considered a sound investment even though it did not achieve perfection in reducing labor turnover to zero. Evaluation of a testing program should be made in terms of a statistical comparison between the employment situation with the tests and the employment situation without the tests. This statistical comparison should include such factors as the productivity of employees, labor turnover, accuracy and safety, ease of shifting employees to different types of work, ease of learning the job, and any other factors that may be indicative of a desirable employee. It seems only fair to evaluate a testing program in terms of averages rather than in terms of specific cases. This point is emphasized because unless one is careful he is likely to allow one serious failure in a testing program to outweigh the less spectacular though much sounder averages that really indicate the value of the method.

Types of Tests. Tests that have been and are being used for employee placement may be classified in several different ways. They may be *group* or *individual* tests. The group variety may be given to almost any number of persons simultaneously, the only limitation on the number usually being seating and writing facilities and provision for adequate hearing of the instructions given by the group examiner before the test is begun. Examples of group tests are the Purdue Achievement Tests (see p. 195), the Adaptability Test (see p. 113), the Purdue Mechanical Adaptability Test (see p. 126), the Otis Self Administering Tests of Mental Ability (see p. 112), the Wonderlic Personnel Test (see p. 113), and the Bennett Test of Mechanical Comprehension (see p. 124). Individual tests, on the other hand, are given to one person at a time and usually call for the undivided, or nearly undivided, attention of the examiner while the test is being administered. The phrase "nearly undivided attention" is used because in certain cases, as with the Purdue Pegboard Test of Manual Dexterity (see p. 138), it is possible for an attentive examiner to test several persons simultaneously if the necessary sets of equipment are available.

Another classification of tests may be made according to whether they are *instrumental* or *paper and pencil.* The former, as the name implies, make use of instruments and ordinarily are individual in nature. The latter make use of written responses and ordinarily are group tests.

A very important division of tests may be made according to whether they measure *aptitude* or *achievement.* (There is still another class of tests that cannot be considered in either of these categories; these are *personality and interest tests,* sometimes called *inventories.*) Aptitude tests measure whether an individual has the capacity or latent ability to learn a given job if he is given adequate training. Such tests are most useful when the majority of applicants for a certain job have had no experience on the job and when a relatively long period of training is required before their aptitude or lack of aptitude for the job will be apparent on the job itself. A good example of a job for which aptitude tests are particularly adapted is looping in a hosiery mill. This is a job that requires a year of experience, on the average, for complete mastery. New employees are usually hired directly out of high school with no previous experience on the operation of a looping machine or anything even remotely resembling it. Aptitude tests that have been found to be effective for this job deal with such factors as finger dexterity (see p. 140) and certain visual characteristics (see p. 146). Interestingly enough, the tests that "came through"[4] in no way resemble the actual looping operation; rather they measure certain basic capacities that are necessary for an efficient performance on that job.

Achievement tests, on the other hand, measure how well the individual can do the job or what he knows about it at the time he is tested. Achievement tests are of greatest value when many of the applicants for a job have had, or claim to have had, experience on the same or a similar job in some other organization. For example, a standardized test of knowledge pertaining to a machine shop will quickly reveal how much the applicant really knows about the machinery and operations in use in a standard shop. It is true that he may evidence considerable knowledge on such a test and still be a poor machinist because, for example, he may lack manual skill. But experience shows that if he reveals little or no knowledge on such a test he is exceedingly unlikely to be a good machinist. This type of achievement test, therefore, is of definite value in eliminating from consideration those who lack the basic information necessary for satisfactory performance on the job. Achievement tests, such as the punch press test described on p. 190, may also be instrumental in nature.

Although the foregoing discussion may seem to imply that a given test may always be definitely identified as an aptitude test or an achievement test, this implication is not always valid. There are certain tests that, when they are used in certain ways, may most properly be considered achievement tests, but that, when they are used in other ways, are more properly classified as aptitude tests. For example, a standard typing test, such as the Thurstone Examination in Typing (see p. 194), would definitely be an achievement test if it were used to select applicants for a typing job. This same test, however, might quite properly be considered an aptitude test if it were used as one element of a test battery designed to select teletype trainees, for in this latter case the typing skill the test measures is basic to the successful accomplishment of the trainee. It should be kept in mind that the basic distinction between aptitude tests and achievement tests must be made in terms of the *purpose* the test is intended to serve—not in terms of the content of the test itself.

---

[4] The phrase "come through" means that the test scores are found significantly related to success on the job.

Personality and interest tests (inventories) are intended to measure personality characteristics or patterns of interests of individuals, on the assumption that such characteristics or interests may be related to performance on various kinds of jobs. When such tests are used as employment tests they are typically used in the prediction of job performance, in essentially the same manner as aptitude tests. It should be added here that there are certain serious limitations to the usefulness for employment purposes of presently available personality and interest tests. These limitations will be discussed in a later chapter.

Definition of Terms. The term *capacity* refers to potential or latent ability. One might have high capacity for a certain skill without having any of that skill at the time of testing. Thus, one might have exactly the combination of keen vision at close distances, finger dexterity, and mechanical ability necessary to become a watchmaker or repairer, but, if he has had no training in watchmaking, he might know nothing about the mechanism of a watch. It is always advisable to know how much capacity one has for a certain job before training is given on the job. Ordinarily those with capacity for the job will learn quickly and efficiently; those without capacity for it will not.

The term *ability* refers to developed capacity, or actual knowledge or skill already developed. One with little capacity might, if he has been given a great deal of training, show more ability at the time of testing than one with high capacity who has not been trained; but if the latter is trained he will soon surpass the former. From a long-time viewpoint, therefore, it is more important to know the capacity of applicants before they are placed than to know their ability. For purposes of immediate addition of personnel who will get out some production, the opposite situation is of course true.

A relationship may be inferred between types of tests and the definitions of capacity and ability given above. Aptitude tests ordinarily measure capacity; achievement tests measure ability.

TESTING THE TESTS—THE EXPERIMENTAL APPROACH

It is of utmost importance in any testing program for selection or placement purposes to use only tests that have themselves been tested or "validated." Only when this has been done can the employment manager have reasonable assurance that the tests will be useful for their intended purposes.

There are two general methods of validating the tests that are commonly used in personnel testing. One of these, the "present-employee" method, consists of testing present employees and correlating the test results with an appropriate criterion of job success. The other, the "follow-up" method, consists of testing new employees at the time of hiring, filing the test results, and later determining the relationship between the test results obtained at the time of hiring and the criterion measures of the employees after they have been on the job for a period of time.

Each of these methods has certain advantages and disadvantages. A long-range testing program, however, preferably should make use of both methods if it is to yield maximum benefit.

In addition to these methods of validating tests, other procedures are occasionally used. Brief mention will be made of certain of these methods later on.

The present-employee and follow-up methods involve certain steps, some of which are alike. A discussion of these steps will clarify the similarities and differences.

Present-Employee Method of Testing the Test. This method of testing the test (test validation) involves the use of a group of present employees on the job or jobs in question.

1. Select battery of experimental tests. An early step in a test-validation project is that of selecting a battery of tests to be tried out. These tests should be chosen in terms of the extent to which they are considered to measure attributes that are judged to be important to job success. This selection preferably should be made on the basis of information obtained from a job analysis.

2. Administer tests to present employees. The tests selected are then given to the present employees on the job. This is usually done on a voluntary basis, in order not to create too much unrest among the employees.

When a plant begins, for the first time, a project of test validation that involves testing a large group of present employees, the question of the reaction of the employees naturally arises. Will they become anxious or unduly excited? Is there a possibility of the testing program causing worker unrest? In this book will be summarized research work and test results that have been obtained by testing hundreds of employees in numerous plants. Many of these plants were strongly unionized. In the union plants, several labor organizations have been represented. In no case has any trouble arisen either with employees or union representatives. Wherever the management has signed a union contract, representatives of the union are contacted and the project explained to them before actual testing is begun. Usually a mimeographed slip reading somewhat as follows is given to each employee:

> The Personnel Department is conducting a series of experiments. You have our assurance that this testing is being done to "test the tests" and that the results will *not* be used, now or later, in any way that will affect your standing with the company.
>
> PERSONNEL DIRECTOR

It should be emphasized that in using this method of testing the test, management should be utterly sincere in its plan to use the results *only* to test the tests. Under no circumstances should the test results be used as a basis for transfer or layoff, or for any other purpose detrimental to the employee or which he may *believe* to be detrimental. After a testing program has been thoroughly tried out and installed, many uses for the test scores may be made, uses that will be advantageous to both employees and management; but in the "testing the tests" stage of the program, management should adhere scrupulously to its agreement not to use the results for any other purpose than the one stated.

3. Select appropriate criterion. At some early phase of a test-validation project it is necessary to determine what criterion of job performance to use. The criterion was discussed in Chapter 2, so we need not repeat that discussion here. The point will be re-emphasized here, however, that the criterion or criteria used in any test-validation study should be *relevant,* meaning that it should reflect the standards by which the performance of employees should be evaluated in terms of management's objectives.

4. Obtain criterion information on present employees. After determining what criterion or criteria to use, it is then in order to obtain criterion information on the individual employees now on the job. Depending on the criterion in question, this may involve the accumulation of available records (such as production records, sales volume, and the like), or it may consist of obtaining ratings from supervisors on job performance of the employees, or other appropriate processes.

Depending on the manner in which the results are to be analyzed, the criterion information may be used to divide the total employee group into two (or possibly three) groups such as "high" (or above average) and "low" (or below average), or it may be expressed in quantitative terms such as units of production or numerical ratings.

5. *Analyze results.* After the test scores and criterion information have been obtained, the results may be analyzed in several different ways. One method is to divide the total group of employees into a "high" criterion group and a "low" criterion group as mentioned above. Next, the employees are divided on the basis of test scores into two, or three, or four test-score groups. For each test-score group the per cent of "high" criterion employees is then computed. If statistically significant differences in these percentages are found for the different test-score groups, it may be safely assumed that the test is measuring something of importance on the job. The results of this type of analysis usually are shown in an "expectancy chart" such as illustrated in Fig. 4.1. This type of chart shows the "odds" of being successful of the applicants within each test-score group.

Another, and even more effective, method of determining the relationship between test results and job performance is to compute a coefficient of correlation between the test scores of the employees and their criterion values. This has certain advantages over the expectancy-chart method mentioned above. In the first place, it gives a more accurate indication of the *amount* of the relationship between test scores and efficiency. In the second place, it enables the employ-

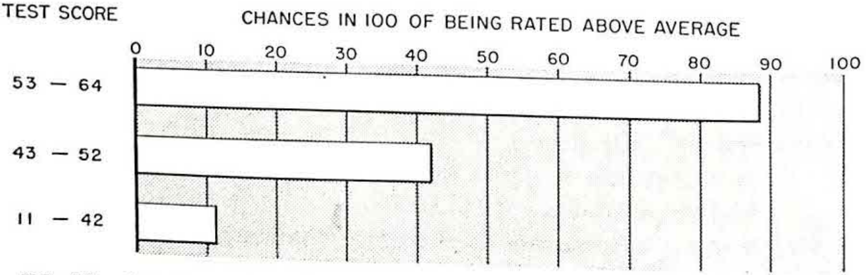


FIG. 4.1. Expectancy chart showing relation between scores made on the Minnesota Paper Form Board (see page 127) and rated success of junior draftsmen in a steel company.

ment manager more effectively to take advantage of the all-important selection ratio (see p. 96) in using the test. In the third place, it makes possible the computation of the relative importance of several tests in an employment battery so that the tests may be "weighted" according to their importance. Finally, the use of the correlational method makes it possible to offset, statistically, whatever influence such factors as experience on the job, or age, may have had on the test scores of the employees. Further mention will be made later of these possible influences.

Although the correlational approach has certain advantages over the expectancy-chart method, these are largely statistical advantages. In terms of understanding by nonstatisticians, the expectancy-chart method probably cannot be excelled.

Follow-up Method of Testing the Test. As indicated earlier, this method consists of giving the test, at the time of employment, to a large number of employees whose placement has already been decided upon by previous, non-test methods of selection and placement. These employees usually are hired just as they have always been hired, but before they have been placed in their jobs, or, in fact, even before they have been told they have been selected for employment, they are required to take the battery of experimental tests. These tests at this time have no effect whatever upon whether the applicants will be hired. The steps involved in this method are described below.

1. Select battery of experimental tests. This step is essentially the same as with the present-employee method.

2. Administer tests to applicants. The tests are administered to applicants who are to be employed for the job in question, but the applicants should not know at the time that a decision has been made to employ them. The test results are then filed until a later date.

3. Select appropriate criterion. This determination is made in the same way as with the present-employee method.

4. Obtain criterion information on the new employees. The criterion information on the new employees usually should be obtained after sufficient time has elapsed for them to demonstrate their actual abilities on the job. Usually this would be after completion of training, or at least after the completion of most of the training.

5. Analyze results. This step is carried out in the same manner as with the present-employee method.

Comparison of the Two Methods of Testing the Test. Each of these methods of test validation has certain advantages and disadvantages. For example, from the logical viewpoint, one objection to the present-employee method is that a test may be measuring something that is improved significantly by experience on the job. In other words, the test in question may be measuring achievement more than aptitude. Consider again the operation of looping in a hosiery mill. If a test is proposed that imitates the operation of a looping machine, it is almost certain that present employees will perform on this test in much the same way that they do the actual job. This means that the test, being in itself a miniature of the job, will simply divide the employees more or less as they have already been divided on the basis of the job itself. Such a test would have little value for placement of new employees unless the surplus labor supply included a large number of employees who claimed to have had looping experience in some other hosiery mill. Under ordinary circumstances, however, persons employed for this job are freshly out of school. They have had no experience in looping, and their capacity for the job would not be measured effectively by a test of this type. Since none of them can loop at the time of employment, all of them, even those who are potentially the best employees if given the training, would score zero or close to zero on any test that calls for ability to loop at the time of employment.

In testing aptitude tests according to this principle it is therefore necessary to be sure that the tests, besides dividing the employees according to their ability on the job, do not show a significant correlation with experience on the job. In other words, it is necessary to be sure that the employees who score high on the test are not scoring high simply because they have had an opportunity to learn the skill being tested. Whether this is the case may be determined by correlating the test scores with experience on the job. If the test is to be used later for employment and/or placement it must satisfy two requirements. First, it must show a positive correlation with ability on the job. Second, the scores on it must not be appreciably related to experience on the job. When these two conditions are satisfied, it may be assumed that the test is *not* measuring something that is improved markedly by experience on the job, but that it *is* measuring something that is necessary for adequate performance on the job. This reasoning applies only to *apitude* tests, not to *achievement* tests.

When tests of the latter type are used, a significant correlation between test scores and experience is to be expected and in no way reduces the value of the tests.

If the test scores show *some* correlation with experience, the net relationship between test scores and job performance, after the effect of experience has been eliminated, can be determined by partial correlation. The procedure for computing partial correlations may be found in any standard textbook of statistics.[5]

Although the influence of experience on test scores in such a situation can be eliminated by the use of partial correlations, a more straightforward manner of getting around such a problem is through the use of the follow-up method. In this case the test scores are obtained *prior* to any experience on the job, and therefore cannot be influenced by job experience.

The foregoing discussion applies only in cases where the applicants have not had experience on the job for which they are being hired. If they have had experience, then an achievement test that is similar to the job or even a *miniature* (that is, the job itself in standardized test form) of the job may identify the best employees even more satisfactorily than an aptitude test.

The present-employee method has certain other possible disadvantages. In some instances, the *present* employees on a job may represent a highly select group, since most of those who were *not* satisfactory either were separated or left of their own accord. In such a case, the correlation between test scores and criterion values would *not* represent the relationship that one would expect with the follow-up method.

Further, the "mental set" of present employees toward taking tests on a voluntary, "experimental" basis may be different from the "mental set" of applicants. This difference in set can influence performance on some types of tests, especially personality and interest tests. Where this influence is of some consequence, it would be preferable to validate the test by the follow-up method. If this is done, the test is then validated in the same type of situation as the one in which it will later be used.

Another disadvantage of the present-employee method is that the arrangements for testing present employees sometimes are difficult to work out, especially since it is necessary to take people away from their jobs in order to test them.

---

[5] J. P. Guilford, *Fundamental Statistics in Psychology and Education* (3rd ed.; New York, McGraw-Hill Book Co., 1956).

Where any of the above conditions exist or might have important influences on test results with the present-employee method of test validation, the follow-up method probably would be preferable. The follow-up method, however, requires more time. Frequently, it does not result in any proved tests until the program has been under way for several months or, in some cases, several years. Should the first battery of tests selected for tryout prove entirely unsatisfactory, it is necessary to start the whole procedure over again. This involves the loss of a great deal of time.

Fortunately, in practice the employment manager need not decide upon one or the other of the methods discussed and then limit himself exclusively to that particular method. He may proceed with both methods simultaneously. When this is done it is possible to produce a working battery of tests almost at once (by using the present-employee method of testing the tests) and to obtain still more evidence of the value of these tests as new employees are tested at the time of placement. This statement may seem to contradict itself because it may be argued that if a test is used for placement there will be no variation in the scores on this test among the new employees placed on any given job and, therefore, it will be impossible to correlate the scores with later success on the job. Actually, even though test scores are used as a basis of employment or placement, there will ordinarily be a sufficient spread of scores on the test, even among employees who are hired for a specific job, to make possible the subsequent statistical evaluation of the test in terms of the follow-up method. Though the variability of the group of employees placed on any given job will be smaller than that of all the applicants, nevertheless the variability of all the applicants and the variability of the applicants who are placed on the job will be known and it will therefore be possible to determine on the basis of the follow-up method just how much the test in question increases the efficiency of the placement procedure for the particular job.

The point to be emphasized throughout this discussion is that no one—whether he is an employment manager, a psychologist, or anyone else—can predict with certainty which tests will be desirable tests for placement on any particular job. Often a rather accurate estimate can be made; but it is usually desirable to check this estimate against correlations between the actual performance of the employees on the job and the scores that they made on the tests at the time of placement.

Other Methods of Test Validation. Although the present-employee and follow-up methods of the test validation are by all odds the most commonly used procedures, and in general the most defensible methods, there are certain other procedures that can be used.

One of these we might call a "job-group" method. In this method the test scores of employees on each of two or several jobs are compared to ascertain whether there are significant differences. In one such instance,[6] the median scores on an intelligence test were obtained for three levels of supervisory personnel (including 983 individuals). These medians are given below:

Job Group	*Median Test Score*
Supervisors (presumably first-line)	116
Department foremen	122
Divisional superintendents	132

The use of this method is predicated on the assumption that by "natural selection" people in an organization tend to "gravitate" into the kinds of jobs that are reasonably in line with their abilities. If it can be assumed that, in general, most persons on a job are achieving a reasonable degree of success on that job, it is then possible to *compare* the employees on various jobs in order to ascertain what *differences* there are from job to job, as in the above example. It is then possible to select for initial placement, or for promotion, those individuals whose test scores most nearly correspond with the test scores of individuals now on the job.

Another procedure that will be mentioned is that of "synthetic" validity. This term was originally used by Lawshe[7] " . . . to denote the inferring of validity in a specific situation. The concept is similar to that involved when the time study engineer establishes standard times for new operations, purely on an *a priori* basis through the use 'synthetic times' for the various elements constituting the operation."

The procedures involved in synthetic validity studies are complex and time-consuming, and will not be discussed in any detail. Suffice it to say that the procedures depend upon two essential points: (1) the identification, in reasonably objective terms, of work elements or other job characteristics—which is essentially a job-analysis process; and (2) the "validation" of a test *on numerous jobs with a common work element* or *job characteristic*. If, in general, it is found that there is a relationship between *scores on a given test*, and *performance*, on *jobs with a characteristic in common*, it is then possible to use that test in the placement of other individuals for *other* jobs with that *same* characteristic.

---

[6] L. Stockford, "Selection of Supervisory Personnel," *Personnel*, 1947–48, **24**, 186–199. *Note:* Median scores reported are approximate values, having been read from graphs.

[7] C. H. Lawshe, "Employee Selection," *Personnel Psychology*, 1952, **5**, 31–34.

This procedure probably offers greatest potential usefulness where there are not enough present employees or applicants available for more conventional test-validation procedures. At present this general approach is possible on only a limited basis, especially with clerical jobs. In a study of clerical jobs, Lawshe and Steinberg[8] found that this procedure offers reasonable promise. They used the *Job Description Check List of Clerical Operations*[9] to identify the work operations of 262 positions in 12 companies. The people in these positions were given the *Purdue Clerical Adaptability Test*,[10] which results in several subscores. By a procedure that need not be discussed here, various operations had been determined to have "critical" requirements on each of the several subscores on this test, such as "spelling."

In general, it was found that the larger the number of "critical" work operations in a job (such as on spelling), the higher were the scores of individuals on such jobs, as shown in Fig. 4.2. This implies that there is a tendency for employees to "gravitate" to those kinds of jobs for which they have some particular aptitude. In the present illustration, it would be reasonable to place on jobs that have been *objectively* determined to have several operations with "critical" spelling requirements, those persons who have high spelling-test scores.

Although the present status of developments in "synthetic" validity does not yet permit its widespread use, this general approach seems to have reasonable future potentialities.

---

[8] C. H. Lawshe and M. D. Steinberg, "Studies in Synthetic Validity. I. An Exploratory Investigation of Clerical Jobs," *Personnel Psychology*, 1955, **8**, 291–301.

[9] Available from Southworth's Extension Service, West Lafayette, Indiana.

[10] Available for purchase from the University Book Store, 360 State Street, West Lafayette, Indiana.

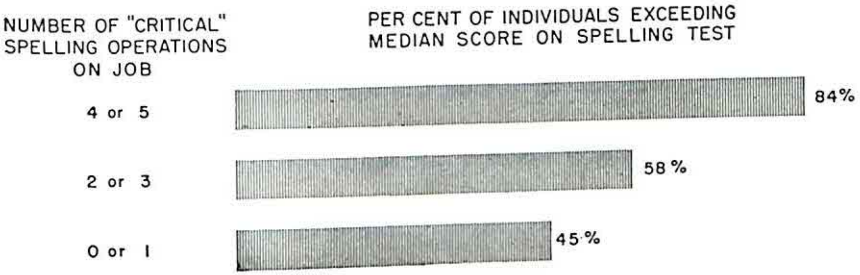


FIG. 4.2. Per cent of individuals exceeding median score on spelling subtest by number of "critical" spelling operations in their jobs. (Adapted from Lawshe and Steinberg, op. cit.)

FACTORS DETERMINING THE FUNCTIONAL VALUE OF PERSONNEL TESTS

There are several factors of both a practical and theoretical nature that determine the functional utility of personnel tests in a business or industrial organization. These include the *validity* and *reliability* of the tests used, the *selection ratio*, and the *percentage of present employees who are "satisfactory" on the job.*

Validity. In a broad sense, the "validity" of a test refers to the degree to which the test is capable of achieving the aims or purposes it is intended to serve. As indicated above, the process of "testing the test" is one of determining the validity of a test for the purposes of selection or placement. We have seen that the *degree* of validity, as reflected by a correlation or by an expectancy chart, affects the *extent* to which the test can be expected to serve its intended purpose of selection or replacement.

Although we are primarily interested in the validity of tests for selection and placement purposes, it may be useful for us to discuss briefly the various specific *types* of test validity that are appropriate in the use of tests in various situations, such as education, vocational counseling, and clinical practice, as well as in personnel work. There are basically four types of validity, these being related to different aims or purposes of testing.[11]

---

[11] A thorough discussion of the concepts of validity and reliability of tests is included in "Technical Recommendations for Psychological Tests and Diagnostic Techniques," *Supplement to the Psychological Bulletin*, March 1954, Vol. 51, No. 2, Part 2, available from the American Psychological Association, 1333 Sixteenth Street, N.W., Washington 6, D.C. Price $1.

Content validity. Content validity is evaluated by showing how well the *content* of a test samples the subject-matter or kinds of situations that it is intended that the test measure. It applies particularly to achievement tests, or tests given in schools. Generally speaking, this determination needs to be made on the basis of the *judgments* of "experts" with regard to the *appropriateness* of the questions or problems (the "content" of the test) as measures of that which the test is supposed to measure.

Predictive validity. Predictive validity is evaluated by showing how well *predictions* made by the test are *confirmed* at some *subsequent time.* The follow-up method of test validation is a typical illustration of this type of validity, since the validity of the test is determined at some time after the test has been given. Usually such validity would be established by determining the relationship between test scores and a criterion of performance, such as performance on a job.

Concurrent validity. Concurrent validity is evaluated by showing how much the scores on the test in question are related to some *present* (concurrent) criterion of performance or status. The present-employee method is an example of this type of validity, since test scores are correlated, or otherwise compared, with a performance criterion. The "job-group" and "synthetic" test-validation methods are also examples, but they involve the use of a "status" criterion rather than a performance criterion.

Construct validity. Construct validity is evaluated by determining what psychological *qualities* a test measures, such as "introversion," or "intelligence." Such validity frequently is determined by correlating one test with another that has been previously found to measure the quality in question, or by the use of a statistical process called "factor analysis" which can identify the extent to which various tests measure a human quality in common.

As indicated above, the present treatment of personnel tests in industry typically involves the concepts of *predictive validity* and *concurrent validity.*

Reliability. The *reliability* of a test is the degree to which a test measures *consistently* whatever it does measure.

There are various ways in which the reliability of tests can be determined. One of these is the *test-retest* method. With this method, a group of subjects are given the same test twice, with a time interval between. The two sets of scores are then correlated. If memory or practice from the first administration can affect significantly the scores of individuals on the second administration, this method is not appropriate. To put it another way, this method is most suitable for use with tests on which memory or practice effects are of limited consequence.

The *alternate-forms* method of determining test reliability is one in which two separate, but equivalent, forms of the test are given to the subjects. Such forms typically include items that are similar in nature, but different in actual content. The scores on these two forms are then correlated to determine the degree of reliability. If the test-retest method cannot be used because of memory or practice effects, the alternate-forms method may be suitable in case there are two alternate forms available.

The *split-halves* method is one in which a test is given to a group of subjects under standard administrative conditions. Each test is scored, however, on each of two halves of the items, and these two sets of scores are then correlated. The two "halves" may be chosen randomly ("chance" halves), or they may consist of alternately allocated items ("odds vs. evens"). This method is most suitable when there are numerous items in the test, and where the test is essentially one of "power" rather than "speed." (A speed test is one in which each item can be accomplished by most subjects, but where speed is important, as in a dexterity test. A power test is one in which the *ability* of the subject to complete the test items is more important than his *speed*.)

It should be added that there are other ways of determining or inferring test reliability, but the above methods are the ones most commonly used. It should also be pointed out that there is no such thing as *the* reliability of a test; rather, reliability is in part a function of the group of individuals tested. In the use of personnel tests, it is desirable to determine the reliability with the individuals on whom the tests are validated. This is not usually done, however. Rather, a person using a test will rely on reliability information based on previous groups of subjects, as reported by the test publisher.

Importance of reliability. One might ask this question: Of what importance is test reliability? This question might be answered by saying that if a test is highly reliable, it is possible to put greater reliance on the scores that individuals receive on the test, than if the test is not very reliable. Since, in most employment situations, one has only a single score on a given test for each applicant, it is useful to know how reliable the test is in order to know how stable such scores are. Generally speaking, the reliability of most tests is considered to be satisfactory if the coefficients of reliability are in the .80's or .90's. There are some circumstances, however, in which tests with lower reliabilities might still be used.

A test may have high reliability and yet be quite worthless for any given employment situation. A test might measure height, weight, or even general intelligence with high reliability and yet show little or no correlation between test scores obtained and success of the employees on the job. In other words, the test might have high *reliability* but low *validity* (predictive or concurrent validity), or even no validity at all. But if a test has a *low* reliability, it is not likely to have a satisfactory validity. This is simply another way of saying that no test is likely to have a higher correlation with anything than that test will have with itself. Thus a test that on repeated testing gives scores that correlate with the first testing only .45 is not likely, except by chance, to relate higher with production or any other criterion of employee desirability. It is primarily for this reason that the psychologist using tests in industry insists that they be reasonably reliable before he makes any attempt to determine their validity for particular situations.

Relationship Between Reliability and Validity. A question that is often raised when the validity or reliability of tests is under discussion is: How reliable and valid must a test be in order to be worth while? This is both a reasonable and a natural question when one considers that our measurements of both reliability and validity are usually expressed in the form of a coefficient of correlation that makes possible a variation in the degree of reliability or validity all the way from zero (or even a minus quantity, if we wish to be theoretically exact) to 1.00. If we remember that reliability in a test is necessary because it limits the validity, we may phrase the question more simply by asking: How high must the validity coefficient of a test be for the test to be worth while?

The answer to this question depends upon the use that is being made of the test. The user of tests is nearly always interested in one of two objectives, but is seldom interested in both at the same time.

Either he is interested in making a careful and accurate aptitude analysis of *each person tested*, which is to be used for individual prediction or vocational guidance, or he is interested in selecting from a large group of individuals a smaller group that, *on the average*, will surpass the larger group in some particular respect. In individual counseling work, which deals with vocational aptitude and guidance, the psychologist is interested in the first-named objective. His work will stand or fall on the accuracy of his predictions for individual clients. He therefore has little use for aptitude tests that do not have a validity sufficiently high to justify their use in individual prediction. The exact value the validity of a test should have, to meet this requirement, is not completely agreed upon by all students of the subject. It is uniformly agreed, however, that the higher the validity of the test the better and that *there is no substitute for validity for individual prediction.*

On the other hand, one may be interested in segregating from a large group of persons tested a smaller group that, on the average will surpass the larger group in whatever trait is being tested. This is, in fact, the situation that confronts the employment manager. He is willing to accept, on the basis of tests, a few individuals who will fail on a given job, and to reject (or place upon some other job) a few who, if placed upon that job, would have succeeded, *if on the whole his percentage of successful placements is higher with the tests than it is without them.* In other words, although the employment manager would like to have every swing result in a home run, he usually recognizes that under normal circumstances this is not practical, and is therefore satisfied if he can improve his batting average. Under these circumstances the validity of the tests can be much lower. But one may still ask: How low can it be? A categorical answer to this question can be given, but the full significance of the answer will be clear only after a thorough study of the next section, which deals with the *selection ratio*. The answer is that a test probably will be valuable, no matter how low the coefficient of validity, if it indicates *some* relationship between test scores and the criterion; or, in statistical terms, if there are fewer than five chances in a hundred that the obtained coefficient of validity could have resulted from the operation of chance factors alone. Often this rule will admit tests whose validity is as low as .30 or even lower. The use of tests with such low validity is sufficiently contradictory to much current thought among psychologists to warrant a fairly detailed justification for the conclusions reached above.

The Selection Ratio. Given a personnel test that has a validity coefficient indicating *some* relationship with the criterion, and given more employees to be placed than can be placed on the job in question, the functional value of the test to an employment manager depends upon the ratio of those placed to those tested who are available for placement. This has been referred to as the "selection ratio."[12] An example will clarify the operation of this principle.

If a certain test is given to a large number of employees for whom a criterion of successfulness as employees is available, and if the scattergram of test scores against the criterion is plotted, the points will ordinarily fall into an oval-shaped area somewhat similar to the oval in Fig. 4.3. The higher the coefficient of validity, the narrower will be the oval; and the lower the validity, the more nearly the oval will approach a circle. A validity coefficient of approximately .60 will result in a scattering of scores approximately covering the oval area shown in Fig. 4.3. Now, if employees are placed without regard to

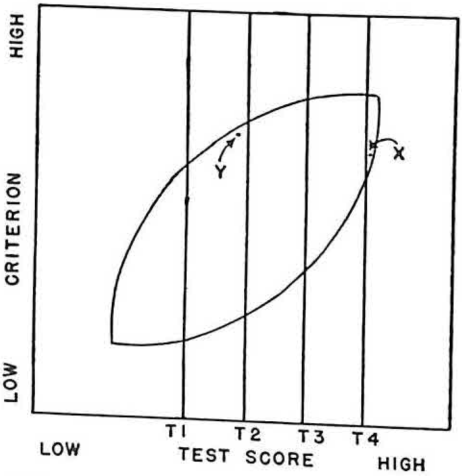


FIG. 4.3. Effect of shifting the critical score required of applicants on average criterion score of employees hired.

---

[12] H. C. Taylor and J. T. Russell, "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables," *Journal of Applied Psychology*, 1939, 23, 565–578.

test scores, their criterion scores usually will be the average of all individuals falling within the oval. If only those are placed upon this job who have test scores as high as or higher than T_1, those not placed on the job will clearly have, on the average, lower criterion scores than the group as a whole, and those placed will accordingly be higher in their criterion scores, on the average, than the group as a whole. A still higher average criterion score for the group placed can be achieved by setting the critical test score at T_2. By moving the critical score to T_3, T_4, or even higher, still more favorable placements, according to average criterion score, can be made.

If a given number of persons, say 60, are to be placed, any one of the conditions mentioned above may exist; which one exists will depend upon the selection ratio that is utilized: that is, the ratio of the number placed to the number tested. Suppose we work with a ratio of 1.00—that is, all those tested are placed. In this case, the distribution of test scores will be over the whole range of possible test scores; the criterion scores will be over the whole range of possible criterion scores; and the test will contribute nothing whatever to the efficiency of the placement procedure. Now suppose that we test 80 individuals and place the 60 who score highest on the test, either not hiring the 20 who score lowest or placing them on some other job. We thus reduce those placed to 75 per cent of those tested, or reduce the selection ratio to .75. Under these conditions, we will place on this job only individuals who test at least as high as T_1, and the average criterion scores of those so placed will clearly be higher than the average of the group as a whole. By testing 120 persons and placing the 60 who score highest on the test, the selection ratio will be reduced to .50 and only individuals to the right of T_2 will be placed. The average criterion score of this group will not only be higher than that of the whole group, but will also be higher than that of the group placed when the critical test score was at T_1. Thus, by increasing the number tested before the 60 to be placed are identified, the selection ratio will be decreased with a continuous increase in the average criterion score for the group of 60 finally placed. If, for example, the plant is expanding so greatly that 600 new employees can be tested before 60 are selected for location upon this particular job (or if the labor market were such that 600 applicants were tested before 60 were hired for this job), the selection ratio would be decreased to .10, only those testing at least as high as T_4 would be placed upon the job, and the average criterion score of the group of 60 placed under these circumstances would be much higher than the score of 60 placed under any larger selection ratio.

The foregoing discussion is of course based on the assumption that the placement of employees is successful in proportion to the average success of the employees placed. Anyone can readily see that even working with a selection ratio of .1, some individuals (like X in Fig. 4.3) will be placed who will be poorer according to the criterion than a few other individuals (like Y in Fig. 4.3) who have not been allocated to this job. But if one is willing to measure the success of the testing program by average results rather than by individual cases, the results will be more and more favorable as the selection ratio is decreased.

We have already stated that psychologists dealing with vocational guidance and individual consultation are usually more interested in making accurate individual predictions than in making group predictions. Most of these psychologists have tended, therefore, to evaluate a test almost entirely in terms of its validity coefficient. They have stressed the fact (which is unquestionably true for individual prediction) that there is no substitute for high validity; that if two tests have been validated against the same criterion, and one has a higher validity coefficient than the other, there is no way to make the one having the lower validity serve as well as the other. The main point of the discussion is that in group testing, where one is interested in average rather than individual results, one can make the test with the lower validity perform as well as the other *by sufficiently reducing the selection ratio.*[13] In other words, in group testing, *a reduction of the selection ratio is a substitute for high validity.* This statement does not mean that this substitute will work if the test has no validity at all, but it does mean that if the test has any significant validity, however small, it is possible for the employer to get the same functional value from it that he could get from a test of any validity, however high, if he is able sufficiently to reduce the selection ratio.

A practical objection to the principle of increasing the efficiency of a test by decreasing the selection ratio may be raised—namely, that there is a limit to the number of applicants who can be tested before the desired number are placed. It is true that, for numerous reasons, an employment manager is seldom able to test 600 or even 200 men before 60 are placed. But it should be remembered that a ratio of 1:10 can be achieved by placing one person out of ten tested just as well as by placing 60 out of 600 tested. One does not need to wait for a great expansion in hiring before advantage can be taken of a reduced selection ratio.

---

[13] *Ibid.* See also Clark L. Hull, *Aptitude Testing* (Yonkers, N.Y., World Book Company, 1928), see footnote on p. 276.

The objection may also be raised that advantage cannot be taken of a reduced selection ratio unless there are more applicants than there are jobs to be filled. This is true if all persons employed are to be placed upon the same job. But almost always an expansion of plant personnel involves hiring for several jobs, not just for a single job. Therefore, the advantage of a reduced selection ratio usually can be achieved even when, as in a period of emergency production, there is difficulty in getting *enough* applicants to fill the jobs. Even when all applicants are hired, placements can be made on various jobs in such a way as to take advantage of individual differences by means of reduced selection ratios. *A reduction in the selection ratio can be utilized whenever two or more employees are being placed on two or more different jobs, if tests of some validity are available for each of the jobs.*

Percentage of Present Employees Considered Satisfactory. A further factor that affects the efficiency of a personnel test in a given employment situation is the percentage of present employees who are considered satisfactory. This factor may be made clear by reference to Fig. 4.4. Suppose we are working with a test having a validity coefficient such that the employees tested fall into the oval-shaped area. Suppose, further, that we are working with a selection ratio of .5—that is, only persons falling in the oval-shaped area to the right of T_2 will be placed on the job. If 50 per cent of the present employees are satisfactory, any increase over this amount in the percentage of satisfactory employees placed that can be achieved by using the test is a gain. Under these conditions, the ratio of satisfactory employees among those placed to the total of those placed would be the ratio of the number of individuals falling to the right of line T_2 and above C_2 to all persons falling to the right of line T_2. This ratio would clearly be higher than .50, and the amount by which it exceeds .50 would be indicative of the functional value of the tests under the conditions discussed.

If all conditions named above remain the same except that previous

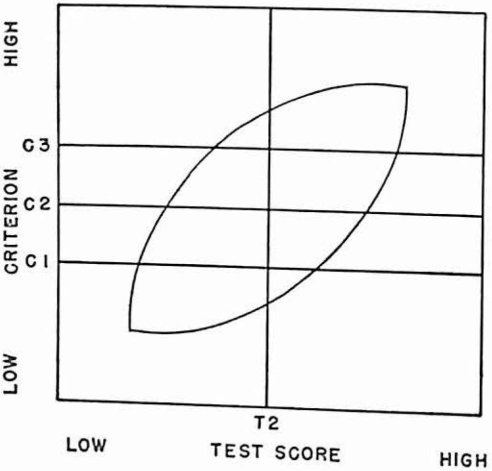


FIG. 4.4. Variation of efficiency of an employment test with differences in percentage of present employees who are considered satisfactory.

employment methods have resulted in 75 per cent satisfactory employees, then the criterion separation line of the successful and unsuccessful employees would be C_1 and the percentage of satisfactory employees placed by means of the test would be the ratio of the individuals to the right of line T_2 and above C_1 to all persons to the right of line T_2. In the latter case, a larger percentage of employees hired will be satisfactory than in the former case, even though the test, selection ratio, and other controlling factors remain the same. In other words if everything else is equal, the smaller the percentage of present employees who have been placed satisfactorily without tests, the larger will be the percentage increase of satisfactory employees when employees are placed by means of test results. This may be illustrated by an example. Suppose we have available a test with a validity coefficient of .50 and are using a selection ratio of .50. Table 4.2 shows the increase, due to using the test, over the percentage of satisfactory employees prior to the use of the test. The values in Table 4.2 were obtained from the Taylor-Russell tables reproduced in Appendix B. If only 5 per cent of employees placed by traditional means are successful, then the expected increase to 9 per cent represents an 80 per cent increase in the number of satisfactory employees placed by the test, under the specified conditions of test validity and selection ratio. For larger percentages of satisfactory employees that have been achieved without the test the percentage of increase achieved by using the test becomes increasingly smaller. If 90 per cent of employees placed by traditional means have been successful, the increase of this percentage to 97 per cent by the test, used under the specified conditions, results in an improvement of only 8 per cent in the number of employees satisfactorily placed.

The general conclusion is that, other things being equal, the more difficult it has been to find and place satisfactory employees without using test procedures, the greater the gain one may expect from a suitable testing program.

The purpose of the foregoing discussion is to make clear that several factors, each relatively independent of the others, operate to determine the functional value of a selection test to an employment manager. If the employment man knows these factors—that is, if he knows the validity of his tests, the selection ratio with which he is working, and the percentage of present employees considered satisfactory—he can predict definitely just how much he will improve the placement process by using the test. And if this amount of improvement is not satisfactory, he can improve his placement by

TABLE 4.2

Increases in Percentage of Satisfactory Employees Placed on a Job over Various Original Percentages of Satisfactory Employees When a Test with Validity Coefficient of .50 Is Used with a Selection Ratio of .50

A	*B*		
Percentage of Satisfactory Employees Placed on the Job *without* the Test	Percentage of Satisfactory Employees Placed on the Job *with* the Test	Difference in Percentage Between Columns *A* and *B*	Percentage of Increase of Values in (*B*) over Values in (*A*)
5	9	4	80
10	17	7	70
20	31	11	55
30	44	14	47
40	56	16	40
50	67	17	34
60	76	16	27
70	84	14	20
80	91	11	14
90	97	7	8

almost any reasonable amount if the selection ratio can be decreased.

Fig. 4.5 reproduces a chart that shows how the percentage of employees selected who will be successful is determined by the validity of the test and the selection ratio. This chart deals with an employment situation in which 50 per cent of present employees are considered satisfactory. The base line in this figure gives the selection ratio, and each of the curves plotted indicates a different test validity. It will be seen that by using a test with a validity of .90, and by reducing the selection ratio to .60, the percentage of satisfactory employees placed will be raised from 50 per cent to 77 per cent. It will also be noted that a corresponding increase to 77 per cent in the number of satisfactory employees will be achieved by a test with a validity of only .50 if the selection ratio is decreased to .20.

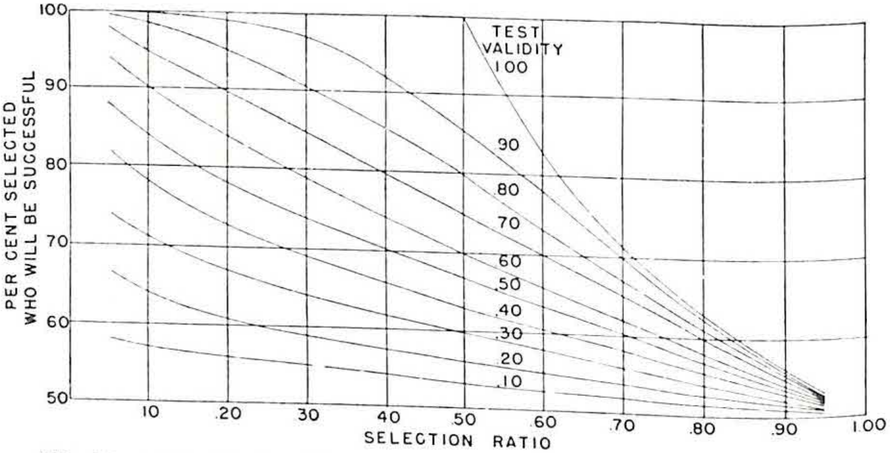


FIG. 4.5. Effect of test validity and the selection ratio upon the working efficiency of an employee selection test.

The Taylor-Russell tables,[14] reproduced in Appendix B, make it possible to determine what percentage of employees hired will be satisfactory under different combinations of test validity, selection ratio, and percentage of present employees considered satisfactory.

The use of these tables may be made clear by an example. Suppose an employment manager has a test with a validity coefficient of .40, has twice as many applicants or employees available as there are jobs to be filled, and is placing in a department where 30 per cent of the present employees are considered satisfactory. Looking in the upper half of the table on p. 550 (entitled "Proportion of present employees considered satisfactory = .30"), we find in the row representing a validity coefficient of .40 and in the column representing a selection ratio of .50, the value .41 where the indicated row and column cross. This means that under the conditions specified, 41 per cent of the employees placed will be satisfactory instead of the 30 per cent attained without the test. If conditions are such that the selection ratio may be still further reduced, the same test will place a still higher percentage of successful employees. For example, if only the highest 10 per cent of the persons tested are placed on the job, the percentage of satisfactory employees will be raised to 58 per cent, or nearly double the percentage of satisfactory employees placed without the test.

---

[14] Taylor and Russell, *op. cit.*

Several cautions concerning the use of the Taylor-Russell tables have been given by Smith.[15] The tabled values do not apply to triangular distributions of test scores plotted against a criterion. Figs. 4.3 and 4.4, which were used to explain the operation of the Taylor-Russell tables, assume that every increase in average test score is associated with an increase in average criterion measure. The tables further assume that the criterion measure in relation to test score is a linear function. Under certain conditions, neither of these assumptions is fulfilled. For example, it is sometimes found that success on a job increases with test scores up to a certain point, but that above this point, further increases in test scores bear no relation or even (in rare cases) a negative relation to job success. Guilford[16] states that an inspection of the scattergram between test scores and criterion is usually sufficient to determine whether the relation is essentially linear. If it seems desirable to test the linearity of the plot, several methods are available.[17] If it is decided that the scattergram represents a definitely nonlinear function, the Taylor-Russell tables should not be used to predict the proportion of successful employees that will be obtained by using the test.

A second caution advanced by Smith concerns the source of the validity coefficient used with the Taylor-Russell tables. When, as is often the case, the validity coefficient of a test is based upon the present-employee method of "testing the test," the validity coefficient obtained will usually be lower than would have been obtained if the follow-up method had been used with a group of new employees. The reason for this difference is that many unsatisfactory employees have terminated their employment and only those doing at least well enough to remain on the job are available for the determination of the validity coefficient of the test. This reduction in the range of the criterion measures reduces the validity coefficient. The difference between the available and the true validity coefficients tends to result in underestimating the increase in satisfactory employees that would be obtained by using the test as a part of employment procedures.

---

[15] M. Smith, "Cautions Concerning the Use of the Taylor-Russell Tables in Employee Selection," *Journal of Applied Psychology*, 1948, 32, 595–600.

[16] Guilford, *op. cit.*, 149.

[17] *Ibid.*, 294.

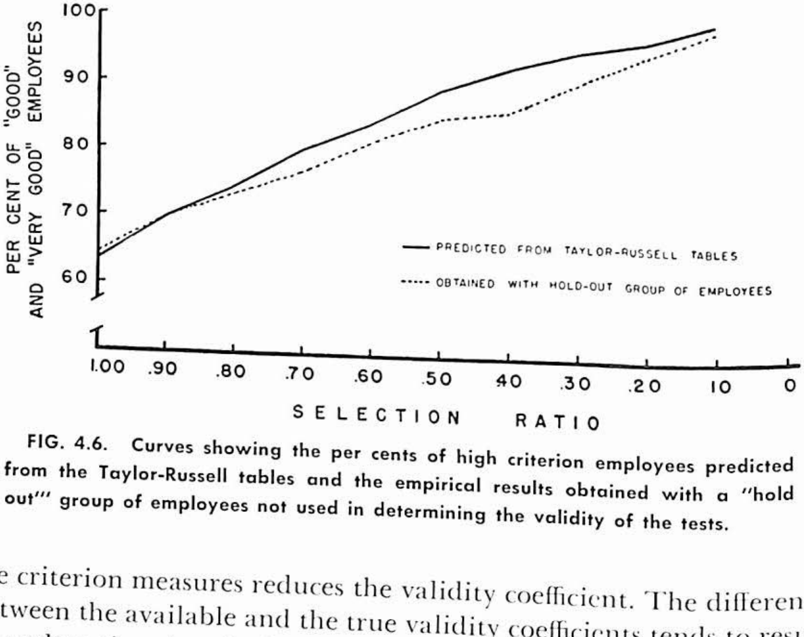


FIG. 4.6. Curves showing the per cents of high criterion employees predicted from the Taylor-Russell tables and the empirical results obtained with a "hold out"' group of employees not used in determining the validity of the tests.

An example of the application of the Taylor-Russell tables to a set of data is shown in Fig. 4.6. The validity coefficient of a battery of three dexterity tests (treated finally as a single measuring instrument as discussed in the next section) was found by Surgent[18] to be .76. A tryout of this battery on a "hold-out" group of employees, who had not been used in determining the validity coefficient, gave the results that have been plotted in Fig. 4.6. Interpolation from the Taylor-Russell tables gave expected theoretical percentages that also have been plotted in Fig. 4.6. The two curves plotted do not differ significantly. It is advisable whenever possible to make this kind of empirical check on the application of the Taylor-Russell tables to personnel test data.

---

[18] L. V. Surgent, "The Use of Aptitude Tests in the Selection of Radio Tube Mounters," *Psychological Monographs*, 1947, 61, No. 2, 1–40.

COMBINING TESTS INTO A BATTERY

No single test will measure all of the capacities or abilities required on any job. Even the simplest of jobs is complex if one considers the combination of capacities or abilities required of a person who is to remain on the job and to do it well. Johnson[19] has pointed out that aptitude for any job consists of a syndrome of abilities and that one needs all of these to be successful. This fact makes it desirable, and in some cases necessary, to use a battery of tests rather than a single test. By statistical methods the results of several tests can be combined into a composite score so that each is weighted to give the maximum correlation between the battery test score and the criterion. An example of the use of such combinations and weighting may be found in a set of tests worked out for placing menders in a hosiery mill. A number of tests were given to 100 employees on this job. For each employee, data were obtained on age and experience as well as average hourly earnings for the twelve-week period preceding the administration of the tests. The correlations between several of the tests and the earnings criterion are given in Table 4.3.

It will be noted in Table 4.3 that the maximum correlation of any individual test with earnings was .27, but that a battery made up of all three correlated .35 with the same criterion. In obtaining the composite score the following formula was used: Composite Test Score = 12 (Hayes Pegboard) − 4 (Purdue Hand Precision Test) − 2 (Finger Dexterity Error Score). The constants by which the raw test scores are multiplied in this formula result in the best combination of the tests to make a prediction of job success from a combination of test scores. The usefulness of a test battery that correlates to the extent of .35 with a criterion may be inferred from Fig. 4.7, which shows the percentage of employees placed by this battery who will be above the average of present employees when

---

[19] H. M. Johnson, "Some Neglected Principles of Aptitude Testing," *American Journal of Psychology*, 1935, **47**, 159–165.

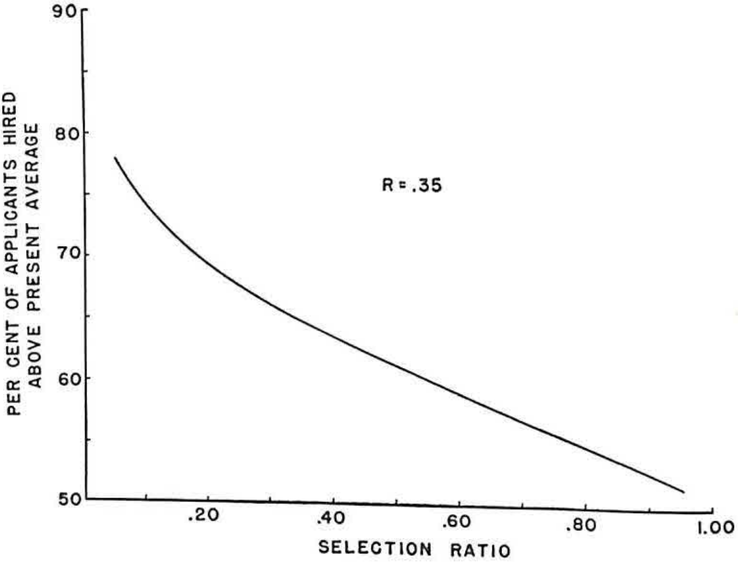


FIG. 4.7. Variation in per cent of employees who will be above the present average of employees when different selection ratios are used with a test having a validity coefficient of .35.

different selection ratios are used in hiring or placement. For example, if the selection ratio could be reduced to .10, approximately 74 per cent of the employees placed would be above the present average. The value of the testing procedure in this case is definitely enhanced by the use of several tests in combination, since no single test, of those tried, gave as high a validity coefficient as did the battery as a whole.

TABLE 4.3

Correlations Between Several Dexterity Tests and Earnings of Menders in a Hosiery Mill

(Effect of age and experience statistically eliminated)

	Correlation with Earnings
Purdue Hand Precision Test	.27
Finger Dexterity Test (Error Score)	.18
Hayes Pegboard	.16
Composite Score of Three Above Tests	.35

Further need for using a battery of tests, rather than a single test, in many cases arises because any single test will at best cover only one of several qualifications that are necessary on the job. A battery is required in order to obtain a more complete picture. The general appearance of the scattergram obtained by plotting test scores against a criterion often, in itself, indicates that a test, even a good test, covers only one phase of the job requirements. The scattergram often shows that persons testing high on the test may be either high or low on the criterion, whereas those testing low on the test are almost always low on the criterion.

An example of the tendency is shown in Table 4.4. The scattergram in this table shows the relation between the learning cost of 35 loopers in a hosiery mill and their scores on the Purdue Grooved Pegboard Dexterity Test (see p. 140). The minimum wage law now in effect requires that every employee be paid a certain minimum wage regardless of whether his piece-rate earnings reach this amount. A slow learner will thus have to be paid a large amount of "minimum make-up" before she is earning her wage, whereas a fast learner will reach the legal minimum more quickly. Thus one indication, and a rather important one, of the desirability of an employee is the total amount of money that the company must contribute in minimum make-up wages while the employee is learning the job. It will be noted in Table 4.4 that this amount varies from $15, in the case of very rapid learners, to more than $60 in the case of slow learners. The grooved pegboard is scored in time required to fill the board (best two trials out of four). The shorter the time, the better the score.

Table 4.4 makes several points clear. Employees with a favorable test score definitely tend to show a lower *average* learning cost to the company than do employees with a poor test score. In this case the correlation between test score (speed of filling the grooved pegboard) and economy in learning cost was .48. This figure may therefore be taken as the validity coefficient of the test for the job in question according to the learning-cost criterion.

But another characteristic of the scattergram shown in Table 4.4 that is commonly found in data of this sort is the *triangular* rather than the *oval*-shaped area into which the employees fall. Several employees who tested high actually turned out to be slow learners. No individuals with poor scores, however, turned out to be rapid

TABLE 4.4

Scattergram Showing Relation between Learning Cost to Company of 35 Loopers in a Hosiery Mill and Scores at Time of Employment on the Purdue Grooved Pegboard Dexterity Test

LEARNING COST TO COMPANY	SCORE ON GROOVED PEGBOARD DEXTERITY TEST				
	60 seconds or over	55–59	50–54	45–49	40–44
$15–$24			2	2	5
$25–$34		1	3	3	
$35–$44			2		3
$45–$59	1			3	
$60 or over	2	3	3	1	1

learners. The significance of this finding is that the test in question apparently is measuring one essential requirement of a rapid learner, but only one requirement. An employee who lacks this requirement is practically certain to be a slow learner. But an employee who tests high, that is, who has an abundance of this requirement, may still be a slow learner if he lacks certain other basic requirements for the job. The triangular shape of the scattergram thus indicates the existence of a hierarchy or syndrome of capacities required for this job.

For the reasons discussed on p. 103, the Taylor-Russell table predictions are not applicable to data of the type shown in Table 4.4. Although reducing the selection ratio in this situation will increase the proportion of rapid learners, the tabled expectancies will not be correct because the scattergram is triangular.

ADEQUATE TRAINING OF TESTERS

The impression that the tester makes upon applicants or employees is of vital importance to the success of a testing program. This fact, obvious to the experienced employment manager, has often been underemphasized or even completely overlooked by psychologists themselves. "Everyone has shoes but the shoemaker's wife," says an old proverb. The professional psychologist often becomes so absorbed in the statistical phases of a test program that he forgets the importance of personal and industrial relations in actual test administration. Often the test administrator spends more time with an applicant than anyone else in the employment office. First impressions are lasting ones, and applicants who are employed may long remember the impressions of the company formed the first day at the plant. And the rejected applicants—upon whose reactions the industrial relations of the company within the community are also dependent—have no further opportunity to acquire a different opinion of the company. The man who administers the tests should be understanding, sympathetic, and courteous. He should give the applicant a feeling of importance. If the test results do not justify employment the applicant should be made to understand that, although he is not properly adapted or trained for the job now open, he may be quite qualified for some other job at a later time or for some job now open in another plant. The applicant should be given an understanding of the fact that he himself would not profit from employment on a job on which he would be likely to fail.

To explain such matters to rejected applicants is neither easy nor routine. It must be done individually, thoroughly, and sincerely. Many of the qualities of a successful salesman are of real value to the test administrator in industry. Testing human beings is not the same as testing materials or processes. People *react*, either favorably or unfavorably, to a test situation. It is one job of the tester—indeed, one of his most important jobs—to be sure that they react favorably. Others may have prepared adequate tests; light, roomy, and attractive rooms may be available in which to give the tests; but unless the test administrator creates a favorable reaction to the testing program, the full advantage of this tool for employee selection and placement is not likely to be realized.

5

APTITUDE TESTS

Tests that are used for prediction measure basic human characteristics or abilities that are related to the *capacity* to develop proficiency on specific jobs. These basic characteristics and abilities can be thought of as aptitudes. In this chapter we will be concerned with tests that measure aptitudes. Various aptitude tests will be mentioned or illustrated, along with a few examples of how aptitude tests are being used in personnel selection and placement.

KINDS OF APTITUDES

Psychologists have done a great deal of research in attempting to determine what the basic human characteristics are, and to develop tests to measure them. Much remains to be learned about such characteristics, but commendable progress has been made in their identification and measurement.

The characteristics we call "aptitudes" can be grouped into five classes, as follows: (1) mental abilities, (2) mechanical and related abilities, (3) psychomotor abilities, (4) visual skills, and (5) a class of other specialized aptitudes. It will be useful to discuss these classes separately.

MENTAL ABILITY TESTS

The term *mental ability* is used synonymously with the term *intelligence*. The field of mental ability testing has been more thoroughly explored than any other area of testing. As a result, we know the occupational fields in which mental ability tests are most likely to be of potential value. (The attention that mental ability testing has received, however, does not indicate that it is necessarily any more important than other types of testing.)

The Nature of Mental Abilities. Although we frequently speak of mental *ability*, we should more properly use the expression mental *abilities*, since it has been found that there are various types of mental abilities. The most comprehensive studies of the nature of mental abilities have been carried out by Thurstone.[1] By a series of factor analyses[2] of mental tests he has identified various primary mental abilities of people. These include the following:

Verbal comprehension.
Word fluency.
Memory.
Inductive reasoning.
Number facility.
Speed of perception.
Spatial visualization.

Although the specific natures of some mental abilities are not yet clear, and though various studies of such abilities are not all in complete agreement, there is nonetheless strong evidence, such as that mentioned above, to suggest that there are several different kinds of mental abilities. These factors are correlated with each other, although the correlations among some of them are much higher than among others. The speed-of-perception and spatial-visualization factors will be discussed later, under "specialized aptitudes" and "mechanical ability" respectively.

Thurstone's studies suggest that each primary mental ability can be measured by a test specifically designed for that ability. Accordingly, such tests have been developed. On the other hand, there are some mental ability tests that include questions or items of different types, resulting therefore in a single score based on the composite of the types of items included in the test. Such tests typically include verbal, numerical, and reasoning questions. Tests of individual mental abilities have been found to be useful in many circumstances; and mental ability tests of the "omnibus" or composite nature have also demonstrated their practicality on a wide basis.

---

[1] L. L. Thurstone, *Primary Mental Abilities*, The Psychometric Laboratory, The University of Chicago, No. 50, September 1948.

[2] *Factor analysis* is a statistical procedure used to identify basic "factors" in sets of data such as tests.

Although there are certain "individual" mental ability tests available, these are used mostly for counseling and clinical purposes. Most of the mental ability tests used for personnel purposes are group tests of a paper-and-pencil nature, so our interest is largely with such tests.

Typical Mental Ability Tests.

Otis Tests.[3] The Otis Self-Administering Tests of Mental Ability are typical of the more widely used and thoroughly standardized tests in this field. This series is called "tests" rather than "test" because it consists of four equivalent forms of a higher examination, designed for high-school students and college freshmen, and an additional four equivalent forms of an intermediate examination, designed for grades four to nine. These examinations are modeled after a group test of mental ability designed by Otis for use in a large commercial establishment. They are made up of logical and arithmetical problems, beginning at a very easy level, such as:

> Which one of the five words below means the opposite of north?
> 1. pole, 2. equator, 3. south, 4. east, 5. west. ( )

The number of the correct answer is placed in the parentheses at the right. Among the special features that the series embodies are: the principle of self-administration, which eliminates the need for a trained examiner; a simplified scoring system; a variety of test material; separate norms for a twenty- or thirty-minute time limit; and a simplified chart for computing "percentile scores" from raw scores on the test.

A person's percentile score[4] indicates that percentage of persons in the group on whom the test was standardized who are at or below the score of the person tested. Thus if one does better on a certain test than 20 per cent of the standardizing group, he is said to have a percentile score of 20 or to be at the twentieth percentile. If he excels 90 per cent of the standardizing group he is said to have a percentile score of 90.

The Otis tests also provide for conversion of raw scores into equivalent I.Q. (intelligence quotient) values. The I.Q. is a measurement that is used widely in school testing of young children, and has become a somewhat common term in popular writing. Since, however, it is used very little in industrial testing programs (because it is not a particularly serviceable measurement when used with scores of adults, and because of a certain feeling of resentment that it arouses in many applicants and employees), we shall not explain its technical derivation in this discussion.[5]

---

[3] Arthur S. Otis, *Otis Self-Administering Tests of Mental Ability* (Yonkers, N.Y., World Book Co., 1922).

[4] The computation of percentile scores is discussed in Appendix A, p. 532.

The Otis tests are in common use in numerous industrial and business employment offices, and a number of investigations that have made use of these tests will be summarized.

Wonderlic Personnel Test.[6] This test is an adaptation of the higher form of the Otis Self-Administering Test; it is particularly adapted to the needs of business and industry. The word *personnel* rather than *mental ability* or *intelligence* is used in the title to avoid the negative reaction that many applicants and employees have to a test that obviously deals with mental ability. The adaptation consisted in selecting from the original Otis Test those items that were found to differentiate most markedly superior from inferior employees on various types of industrial and business jobs. The result of this careful selection of items is that for many industrial jobs the revision differentiates more satisfactorily than does the original form, and that the test may be given in twelve minutes instead of the twenty or thirty required for the original test. The general form of the questions on the Wonderlic test is similar to that of the Otis test. The test is begun with such items as:

> Of the five things below, four are alike in a certain way. Which is the one not like these four?
> 1. smuggle, 2. steal, 3. bribe, 4. cheat, 5. sell. ( )

The correct answer of course is *sell*, the other four choices all being alike in the fact that each involves an element of dishonesty.

The Adaptability Test. The Adaptability Test[7] measures mental adaptability or mental alertness. The test has a fifteen-minute time limit and is useful in helping to identify persons who should be placed on jobs that require rapid learning and/or the development of independent judgment. It also aids in identifying persons who do not readily adapt to new situations but who would be satisfactory (and often superior) employees on simple, routine jobs such as packing, inspecting, assembling, or in the operation of simple, repetitive machines.

---

[5] The interested student will find a description of the meaning and derivation of the I.Q. in any elementary textbook of psychology, such as E. J. Asher, J. Tiffin, and F. B. Knight, *Introduction to General Psychology* (Boston, D.C. Heath & Company, 1953).

[6] E. F. Wonderlic, *Personnel Test, Form D* (New York, Psychological Corporation).

[7] Joseph Tiffin and C. H. Lawshe, Jr., *The Adaptability Test* (Chicago, Science Research Associates).

The test was constructed specifically for industrial use. The questions were phrased in practical rather than academic terminology.

The introductory paragraph on the title page of the test was developed through consultation with industrial personnel men and actual try-out with applicants and employees. This introduction, which is reproduced below, makes no mention of mental ability, intelligence, or the I.Q., and represents an effort to put the individual taking the test at ease.

> Some jobs require figuring—such as adding, subtracting, multiplying, and dividing—while others require writing reports or answering letters, and still others can be done well by people who are not particularly apt with figures or words. This test will help in determining how well you can handle jobs that require these abilities.
>
> Do as well as you can on this test, but do not worry about it. Remember that you may be well qualified for certain jobs that require training or skills different from those covered in this test.
>
> (*Reproduced with permission of Science Research Associates*)

The Purdue Non-Language Personnel Test.[8] This is a group test of mental ability that is strictly nonverbal, i.e., it does not call for the ability to read. Since it consists entirely of geometric forms it is also essentially culture free. People with different foreign language backgrounds make scores that are determined by their basic mental ability no matter what the language they speak, read, or write.

SRA Primary Mental Abilities Tests.[9] This series of tests measures some of the primary mental abilities that have been identified by Thurstone. In particular, the series includes the following tests that may in our present context be considered to be mental tests:

---

[8] Joseph Tiffin, Alin Gruber, and Kay Inaba, *The Purdue Non-Language Personnel Test* (Chicago, Science Research Associates).

[9] Distributed by Science Research Associates.

Verbal-Meaning.
Quantitative.
Number.
Reasoning.

(The series also includes a Motor test of eye-hand coordination, a Perception test, and a Space test. The abilities measured by these tests will be mentioned in later sections of this chapter.)

These tests are used primarily for testing school children, there being three different levels for three age groups.

Flanagan Aptitude Classification Tests.[10] This series of tests (abbreviated as the FACT tests) has been developed primarily for personnel purposes (in contrast with the SRA tests above which are for testing children). This series also includes tests of certain "mental abilities" as well as of other characteristics, including the following:

1. Inspection (ability to spot flaws or imperfections).
2. Coding (speed and accuracy of coding office information).
3. Memory (ability to remember codes).
4. Precision (speed and accuracy of finger movements).
5. Assembly (ability to "see" how objects would fit together).
6. Scales (speed and accuracy of reading scales, graphs, and charts).
7. Coordination (arm and hand movements).
8. Judgment and Comprehension (ability to reason logically and to use good judgment).
9. Arithmetic (skill in working with numbers).
10. Patterns (ability to reproduce patterns accurately).
11. Components (ability to identify component parts, as in drawings).
12. Tables (reading tables of data).
13. Mechanics (understanding of mechanical principles).
14. Expression (feeling for, and knowledge of, English).

Of the above, the tests that measure "mental abilities" in our present context are Memory (3), Judgment and Comprehension (8), Arithmetic (9), and Expression (14). The others measure special abilities or aptitudes of other types. The complete series is intended to measure most of the various aptitudes required in the gamut of industrial jobs.

---

[10] Distributed by Science Research Associates.

Sources of Information on Mental Ability Tests. The preceding tests have been mentioned only to illustrate the kinds of mental ability tests that are available. It is not our purpose here to attempt even a partial listing, to say nothing of an evaluation, of the many excellent measuring devices available in this field. *The Mental Measurements Yearbook*[11] gives a very complete listing of intelligence tests. For most of the tests listed, the *Yearbook* gives the title, a description of the group for which the test was constructed, the date of copyright or publication, whether the test is individual or group, the number of forms, the cost, the time required for administration, the author and the publisher, references to studies dealing with validity and reliability, and a brief evaluation of the test by one or more competent authorities. The industrial personnel man looking for a test that will serve some particular purpose can usually decide from the information given which test will be most likely to serve his purpose.

Another list of mental ability tests that have been used successfully in business and industry is reproduced in Appendix C on pp. 554–555.

Examples of Mental Ability Tests in Use. Since mental ability tests have been used with a wide variety of jobs, only a few typical examples of their use in personnel selection and placement will be given here.

Clerical jobs. This is one of the job areas in which mental ability tests have proved to be especially useful. It must be recognized, however, that clerical jobs vary greatly in their mental requirements. Where performance on a job depends predominantly upon other aptitudes or abilities, a mental ability test shows little correlation with job success.

Typical of studies in this field is one made by McMurry,[12] who found correlations varying from .34 to .57 between Otis Intelligence Test scores and job efficiency among bank employees. The employees considered were for the most part engaged at least partially in a type of clerical work that involved careful and accurate computation and machine operations. Pond and Bills[13] have reported a study that further shows the importance of intelligence tests in selecting individuals for certain types of clerical jobs. Their investigation shows that on the easiest jobs the greatest turnover is among the employees with highest test scores, whereas on the more difficult jobs the greatest turnover is among the employees testing lowest in mental ability. In other words, the more difficult clerical jobs call for employees with the type of ability that a mental ability test measures. If all the employees possessing such ability are placed upon such jobs, whereas those not possessing such ability are placed on the easier jobs, the total turnover of the employees on both types of jobs will be definitely reduced.

---

[11] Oscar K. Buros, *The Fourth Mental Measurements Yearbook* (New Brunswick, N.J., Rutgers University Press, 1953).

[12] R. N. McMurry, "Efficiency, Work-Satisfaction and Neurotic Tendency. A Study of Bank Employees," *Personnel Journal*, 1932, **11**, 201–210.

[13] M. Pond and M. A. Bills, "Intelligence and Clerical Jobs. Two Studies of Relation of Test Score to Job Held," *Personnel Journal*, 1933, **12**, 41–43.

In the selection of cashiers, Clarke[14] has reported correlations as high as .57 between productive efficiency (as measured by transactions handled per day) and scores on the Otis Test, a Sales Checking Test, a Change Making Test (speed and accuracy), manual dexterity, and the Bernreuter Scale. The Bernreuter Scale, which will be discussed later (see p. 163), could hardly account for all of the relationship found. It seems clear, therefore, that the abilities tapped by the mental ability test are of primary importance in determining the job efficiency of the cashiers studied. A study by Stevens and Wonderlic[15] reports the relationship between scores on the Otis Mental Test and ability to handle office detail. Their study indicates that the employees who missed most items on the Otis Test are the ones most criticized by company supervisors.

Tiffin and Lawshe[16] found correlations ranging from .40 to .65 between rated success of paper mill clerical employees and scores on the Adaptability Test. In one department where only 57 per cent of current employees were considered satisfactory, it was found that among those employees making scores of 25 or above on the Adaptability Test, 86 per cent were satisfactory employees.

Hay[17] reports a correlation of .56 between Otis Test scores and production of machine bookkeepers. In another study, Hay[18] reports a correlation of .35 between rated ability of clerical employees and Otis Test scores.

---

[14] W. V. Clarke, "The Evaluation of Employment Tests," *Personnel*, 1937, **13**, 133–136.

[15] S. N. Stevens and E. F. Wonderlic, "The Relationship of the Number of Questions Missed on the Otis Tests and Ability to Handle Office Detail," *Journal of Applied Psychology*, 1934, **18**, 364–368.

[16] Joseph Tiffin and C. H. Lawshe, Jr., "The Adaptability Test: A Fifteen Minute Mental Alertness Test for Use in Personnel Allocation," *Journal of Applied Psychology*, 1943, **27**, 152–163.

[17] Edward N. Hay, "Predicting Success in Machine Bookkeeping," *Journal of Applied Psychology*, 1943, **27**, 483–493.

Supervisors. Mental ability tests, in general, have "come through" more consistently in the selection of supervisors than any other type of test. As an illustration, a rubber company recently promoted seventy operators to supervisory jobs. To train these men for their new duties and responsibilities, the company conducted a series of supervisory training classes, at an expense to the company of $225 per man for the time spent in attending the classes. This amount did not include instructional costs or other expenses. During the first meeting of the supervisory training class, the men were given the Adaptability Test discussed on p. 113. The test scores were filed and a period of six months was allowed to pass while various criteria of success on the job of the new supervisors were investigated. At the end of the six-month period, it became apparent that the most satisfactory criterion of success available in this instance was whether or not the man was a good enough supervisor to be still on his new job six months later. Approximately one-fourth of the men were not, having quit or been demoted, transferred, or dismissed. In the case of the men no longer on the job, the cost of the supervisory training, as well as the other considerable expenses incurred in breaking in the new men, was borne by the company without return. Fig. 5.1 shows the percentage of men still on the supervisory job six months later among those scoring at various levels on the Adaptability Test.

In addition to remaining longer on supervisory jobs, other studies have shown that men scoring relatively high on a mental ability test do better work on jobs of this type. Lawshe[19] reports a study in which representatives of 44 different companies, widely scattered geographically, each identified two of their "very best first line supervisors" and two of their "poorest first line supervisors." The four men in each plant, 176 in all, were given the Adaptability Test. The results of this study are shown in Fig. 5.2. Lawshe[20] interprets Fig. 5.2 as follows:

---

[18] Edward N. Hay, "Tests in Industry. I. Their Proper Use," *Personnel Journal*, 1941, **20**, 3–9.

[19] C. H. Lawshe, Jr., "How Can We Pick Better Supervisors?" *Personnel Psychology*, 1949, **2**, 69–73.

[20] *Ibid.*

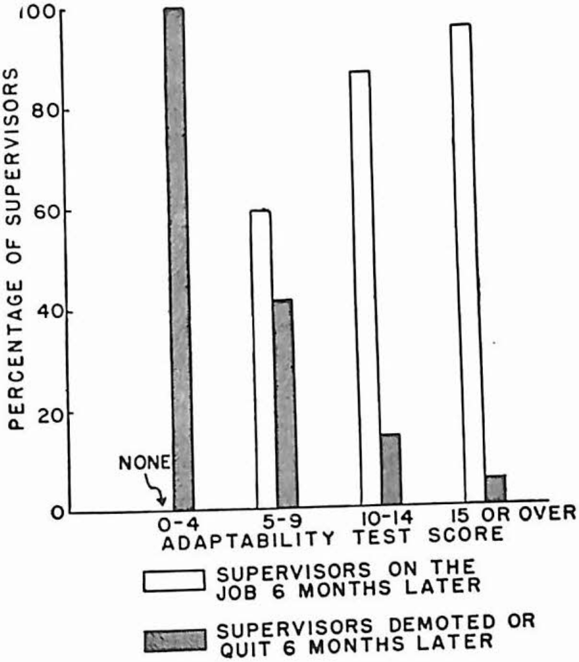


FIG. 5.1. Percentages of men promoted to supervisory jobs in a rubber company who were, and were not, still on the supervisory job six months later.

"Note that when we talk about all of these supervisors, 50 per cent are considered best, since each plant identified equal numbers of poor and best. It seems quite significant, however, that when we direct our attention to the 130 who did best on the test (those making a score of 12 or better) the percentage considered best is 63. Note that as successively higher minimum test scores are established, correspondingly higher percentages of supervisors appear in the best category." Holmes[21] reached similar conclusions in a study of office supervisory personnel, where he found a correlation of .51 between scores on the Wonderlic Test and efficiency ratings of 100 office supervisors. Harrell[22] reports a correlation of .37 between Otis Test scores and ability of cotton mill supervisors.

---

[21] F. J. Holmes, "Validity of Tests for Insurance Office Personnel," *Personnel Psychology*, 1950, 3, 57–69.

[22] T. W. Harrell, "Testing Cotton Mill Supervisors," *Journal of Applied Psychology*, 1940, 24, 31–35.

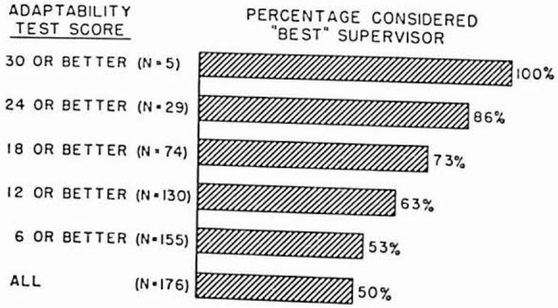


FIG. 5.2. The per cent of employees at or above various test score levels who were considered "best" in 44 industrial plants.

The findings of these studies are in agreement with those reported by Moore,[23] who emphasizes the importance of using some form of mental ability test, in connection with other requirements, in the selection of supervisors.

Other jobs. in a study based on the entire force of a utility plant, Wadsworth[24] reports that intelligence correlated .68 with "man-to-man" estimates by supervisors. He also reports that after testing was adopted as a part of the employment procedure, the percentage of unsatisfactory employees hired dropped from 29 per cent to 5.5 per cent.

Anderson[25] reports that the Adult Placement Test, which measures general intelligence or mental alertness, correlated .40 with efficiency ratings of 174 employees on jobs involving machine skills.

Another area in which the use of a mental ability test has been found very helpful is in the selection of teletype operator trainees. It is customary for the operating company to send trainees for this job to a special school where a carefully prepared program of instruction is given. The trainees are graded periodically by means of standardized and objective tests that cover both job performance and related information. Trainees are certified as operators only after successfully passing this course.

---

[23] H. Moore, "Supervision. I. Selection," *Personnel Journal*, 1942, **20**, 353–356.

[24] G. W. Wadsworth, "Tests Prove Worth to a Utility," *Personnel Journal*, 1935, **14**, 183–187.

[25] R. G. Anderson, "Test Scores and Efficiency Ratings of Machinists," *Journal of Applied Psychology*, 1947, **31**, 377–388.

In a recent investigation, 219 teletype trainees were given the Adaptability Test, previously described, shortly after being selected for training. The follow-up method of test validation was employed to determine the relation between test scores and success in the training course. The results are summarized in Table 5.1.

TABLE 5.1

Relation between Scores on the Adaptability Test and Success in a Training Course for Teletype Operators

Score on Adaptability Test	Number Passing Course	Number Failing Course	Total Number	Per Cent Passing Course
22 or above	27	1	28	96%
16–21	66	4	70	94%
10–15	83	10	93	89%
9 or below	17	11	28	61%

The cost to the company for each trainee (which includes transportation to and from the school, board and room, salary during that training, and the pro-rata cost per trainee of operating the school) amounts to approximately $500 per trainee. The investigation cited above shows that from every 100 trainees who score 22 or above on the Adaptability Test, the company will be able to certify 96 operators, the cost thus being $\frac{100 \times \$500}{96}$ or $520.83 per operator. Among trainees scoring 9 or below on the Adaptability Test, of every 100 trained only 61 will be certified as operators, the cost thus being $\frac{100 \times \$500}{61}$ or $819.67 per operator. The comparative costs show that *each of the operators in the low-score bracket costs the company $298.84, or approximately 57 per cent, more in training expenses than do operators in the high-score bracket.*

There are, of course, many jobs for which mental ability tests have been found *not* to be useful in selection. In such cases, job performance presumably depends upon other aptitudes or abilities, and the degree of mental ability of individual workers neither increases nor decreases the chances of job success.

Still other investigations have shown that mental ability tests sometimes show a negative relationship with success on the job.

For example, Tiffin and Greenly,[26] in studying routine manipulative assembly, found a negative correlation between the Otis scores and both the actual amount of production and supervisors' ratings of productivity. Tiffin and Lawshe[27] found that the best inspectors of simple material made lower average Adaptability Test scores than did the poorest inspectors on the same job. Abel[28] reports a study of 84 girls from the Manhattan High School for the Women's Garment Trades who left school at seventeen years. The average Otis intelligence test score of the girls indicated feeble-mindedness or near feeble-mindedness, and yet after a period of 3.5 years, 55 per cent were steadily employed. Abel points out that among the factors contributing to their success were stable homes, ambition and self-respect, and encouragement and patient treatment during the initial work period.

These studies are in accord with studies already referred to by Bills[29] and by Pond and Bills[30] showing that individuals who test high on mental ability usually should not be placed on routine, easy jobs, and that frequently persons scoring very low on mental ability tests consistently do better on simple, repetitive jobs than do persons whose mental ability is average or above.

MECHANICAL ABILITY AND RELATED TESTS

The term "mechanical ability" has been rather loosely used to characterize a wide range of jobs that involve working with mechanical parts and equipment. We should hasten to say that essentially two *types* of abilities are required, in varying degrees, in "mechanical" jobs. There are certain *mental* aspects, such as mechanical comprehension and the understanding of mechanical principles; and, there are *motor* or *physical* skills such as muscular coordination and dexterity. We will consider mechanical ability in the sense in which it is used by Thurstone[31] as "a complex of intellectual abilities." Although it is true that mechanics, craftsmen, and skilled machine operators do require certain motor skills in order to perform their job, such employees usually succeed or fail in proportion to their training and general mechanical comprehension, not in proportion to their muscular skills. Mechanical ability tests are used primarily to select employees for jobs that require a mechanical "knack," such as in the maintenance of machinery, in "setting up" production machinery before it is turned over to an operator, or in the repair of household appliances.

---

[26] Joseph Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, 1939, 23, 240–263.

[27] Tiffin and Lawshe, *op. cit.*

[28] T. M. Abel, "A Study of a Group of Subnormal Girls Successfully Adjusted in Industry and the Community," *American Journal of Mental Deficiency*, 1940, 45, 66–72.

[29] M. A. Bills, "Relation of Mental Alertness Test Scores to Positions and Permancy in Company," *Journal of Applied Psychology*, 1923, 7, 154–156.

[30] Pond and Bills, *op. cit.*

The Nature of Mechanical Ability. Although we are considering mechanical ability in the restricted sense of mental ability, we still want to know whether there is a single *general* mechanical ability or whether there are *various* such abilities. The accumulating evidence of psychological research seems to justify the conclusion that there are various *kinds* of abilities that may appropriately be considered as mechanical abilities. There is, however, a central or principal mechanical ability that we can consider as something of a general human ability. This was shown, for example, in factor analyses by Friedman and Ivens[32] and Friedman and Detter[33] of certain Air Force and civilian tests. In both of these studies a factor called "mechanical experience" was identified, this representing the knowledge gained from experience with mechanical apparatus and tools. For convenience of terminology, we might call this "general mechanical aptitude."

In addition to this general factor, however, other more specific factors were identified which appropriately may be considered as falling within the area of mechanical abilities. Both studies revealed a factor of *visualization;* this represents the ability to manipulate mentally visual images in two or three dimensions, as in visualizing the forms that various geometric figures would make. This presumably corresponds to the spatial visualization factor reported by Thurstone[34] in his studies of primary mental abilities previously discussed under mental ability tests.

---

[31] L. L. Thurstone, An Analysis of Mechanical Aptitude, The Psychometric Laboratory, The University of Chicago, No. 62, January 1951.

[32] Gabriel Friedman and Frank C. Ivens, "Factor Analysis of the Airman Classification Battery AC–1B; the USES General Aptitude Test Battery, Experimental Visualization and Spatial Tests, and Psychomotor Tests," Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas, *Research Bulletin*, AFPTRC–TR–54–67.

[33] Gabriel Friedman and Howard M. Detter, "Factor Analysis of Airman Classification Battery and Selected Air Force and Civilian Tests," *Research Bulletin*, AFPTRC–TR–54–75.

Another factor reported by Friedman and Ivens[35] was one called *spatial relations*. This may appear to be somewhat akin to the visualization factor, but there is a difference. This factor represents the ability to *perceive* spatial patterns accurately, and to *compare* them with each other, such as in recognizing objects to be identical or different. The visualization factor deals more with the ability to manipulate visual images. It should be added that there is some evidence that there may be other spatial factors[36] in addition to the ones mentioned above.

Although the "general mechanical aptitude" factor probably is the most important in many mechanical jobs, it can readily be seen that these factors of visualization and spatial relations would be of considerable importance in some mechanical jobs, since some such jobs require the ability to visualize the physical relationships of objects, or to identify objects as being similar or different. It should also be pointed out that these various factors are not entirely independent, but tend to be correlated with each other.[37]

Group Tests of Mechanical Ability. By far the majority of mechanical ability tests are group tests, in that they can be given to several people simultaneously.

Bennett Mechanical Comprehension Tests.[38] In these tests the person tested is required to make a judgment based on drawings that are accompanied by simply phrased questions. The problems generally require the application of mechanical or physical principles to practical problems. Two illustrations are shown in Fig. 5.3. There are several forms of this test series, varying largely in difficulty level. These tests measure predominantly general mechanical aptitude.

---

[34] L. L. Thurstone, *Primary Mental Abilities, op. cit.*

[35] Friedman and Ivens, *op. cit.*

[36] L. L. Thurstone, *Mechanical Aptitude. III. Analysis of Group Tests,* The Psychometric Laboratory, The University of Chicago, No. 55, May 1949.

[37] *Ibid.*

[38] These tests, developed by George K. Bennett and his associates, are distributed by the Psychological Corporation. There are various forms, as follows:

Form AA, for general population testing.

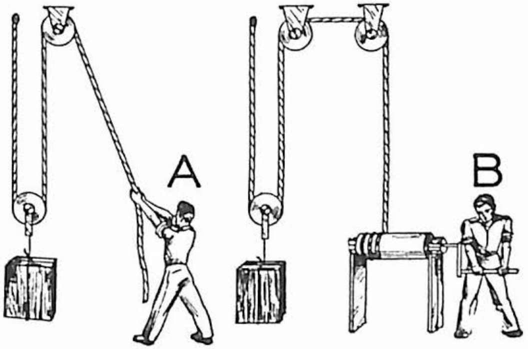
Form AA-F, same as AA except that the instructions and questions are in both English and French, for use with French Canadians.

Form BB, more difficult than AA.

Form CC (Owens-Bennett), more difficult than BB, and yields a wider range of scores at high ability levels.

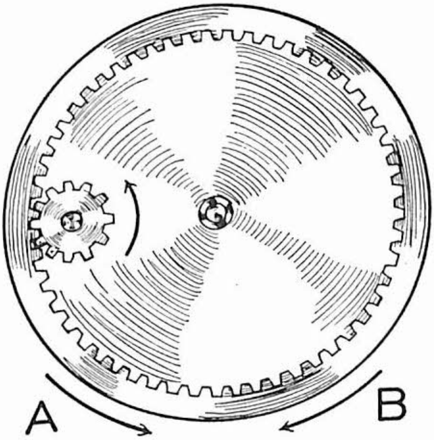
Form W 1 (Bennett-Fry), form for women. Other forms, called the D.A.T. Mechanical Aptitude Tests, have been prepared especially for junior and senior high-school groups.

Flanagan Aptitude Classification Tests.[39] As indicated above,



Which man can lift more weight?

FIG. 5.3. Two items from Form AA of the Bennett Test of Mechanical Comprehension.



If the small wheel goes in the direction shown, in which direction will the large wheel go?

---

[39] Available from Science Research Associates.

these are several tests in the FACT series. Some of these tests are of a mechanical nature. The *Mechanics* test (No. 13) deals with the understanding of mechanical principles, presumably measuring essentially general mechanical aptitude. The *Assembly* test (No. 5) tests the ability to "see" how objects would fit together, this probably being most closely related to the visualization factor.

Purdue Mechanical Adaptability Test.[40] The Purdue Mechanical Adaptability Test, Form A, is designed to aid in identifying men or boys who are mechanically inclined, and who, therefore, are most likely to succeed on jobs or in training programs calling for mechanical abilities and interests. The test measures one's experiential background in mechanical, electrical, and related activities. The test was constructed to measure experiential background because there was reason to believe from a previous study[41] that, other things being equal, those persons who have most profited in knowledge from previous mechanical experiences may do better on mechanical jobs than those persons who have not so profited. The test is one of general mechanical aptitude.

The questions comprising Form A of the Purdue Mechanical Adaptability Test were selected by statistical methods designed to achieve maximum reliability of the final test and as low a correlation as possible with general intelligence. The odd-*vs.*-even-items method of determining the reliability, using a group of 487 industrial applicants,[42] resulted in a reliability coefficient of .84. With the same group of applicants, the correlation between scores on the Purdue Mechanical Adaptability Test and the Adaptability Test was found to be .32. A later study[43] of 382 employees of an electrical manufacturing company showed a correlation of .26 between scores on the Purdue Mechanical Adaptability Test and the Wonderlic Personnel Test.

Industrial Training Classification Test.[44] This test is not strictly a mechanical ability test, but it serves somewhat the same purpose since it is designed to aid in selecting those persons who are most likely to profit from an industrial training program. In particular, it measures the ability to read simple measurements and solve simple arithmetic problems. These simple but basic skills are essential to success in many trade schools and industrial training programs.

---

[40] Available from the University Book Store, West Lafayette, Indiana.

[41] C. H. Lawshe, Jr., and G. R. Thornton, "A Test Battery for Identifying Potentially Successful Naval Electrical Trainees," *Journal of Applied Psychology*, 1943, **27**, 399–406.

[42] C. H. Lawshe, Jr., Irene Semanek, and Joseph Tiffin, "The Purdue Mechanical Adaptability Test," *Journal of Applied Psychology*, 1946, **30**, 442–453.

[43] R. G. Wells, *Industrial Norms for the Purdue Mechanical Adaptability Test* (M. S. thesis, Purdue University, 1950).

[44] Distributed by Science Research Associates.

Revised Minnesota Paper Form Board.[45] In this test the individual chooses which one of five sets of geometric designs can be formed by a given combination of parts. This measures essentially the visualization factor previously mentioned. Fig. 5.4 shows one of the problems.

Individual Tests of Mechanical Ability. To a considerable extent, group tests of mechanical ability can accomplish as much as individual tests.[46] There usually is, of course, a saving in time and cost in using group tests. In some circumstances, however, individual tests offer advantages that justify their use. A few examples of such tests will be discussed below.

Minnesota Mechanical Assembly Test. This test of mechanical ability has been found useful in a number of industrial situations. One form of the test is illustrated in Fig. 5.5. It consists essentially of a number of commonplace mechanical devices, such as a safety razor, a mousetrap, and an electrical plug fixture, all of which may be easily taken apart and re-assembled. The person tested is required to assemble the parts of each article into the finished product. One criticism that has been leveled against this type of test is that the persons tested are likely to vary in their familiarity with the various items, and hence the test is likely to measure this

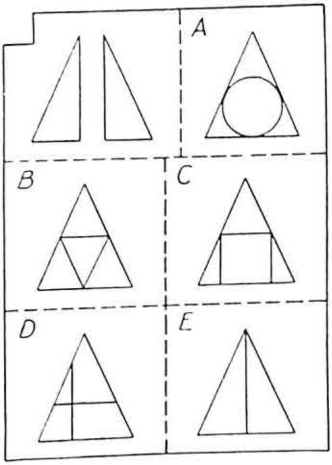


FIG. 5.4. An item from the Minnesota Paper Form Board Test. The person tested is asked to select the set of lettered parts (A, B, C, D, or E) which may be formed by the parts shown in the upper-left square.

---

[45] Distributed by the Psychological Corporation.

[46] W. Harrell, "A Factor Analysis of Mechanical Ability Tests," *Psychometrika*, 1940, 5, 17–33.

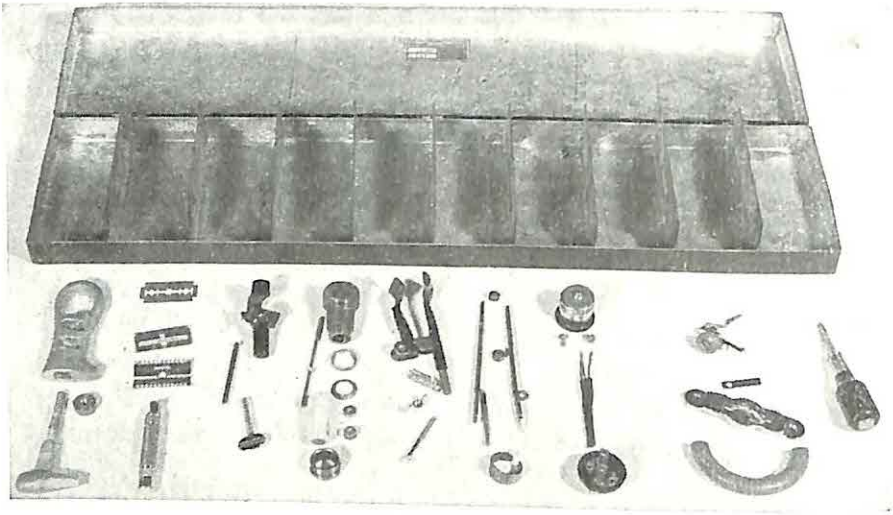


FIG. 5.5. One form of the Minnesota Assembly Test.

familiarity rather than mechanical ability. In spite of the validity of this criticism from a theoretical viewpoint, the fact remains that, in practical situations, the test has shown itself to be a serviceable measuring scale. This fact, rather than a theoretical criticism, is the basis upon which it should be evaluated.

O'Connor Wiggly Block Test.[47] The O'Connor Wiggly Block Test consists of nine pieces of wood cut after the fashion of a three-dimensional jig-saw puzzle so that the parts may be assembled into a single rectangular block. Experimental investigations of this test, such as the one reported by Remmers and Schell,[48] indicate that its validity is high within the limit of its reliability, but that the coefficient of reliability of the test is only .35. This low reliability is no doubt due, at least in part, to the fact that the entire test consists of assembling a single block; that is, the test really consists of only a single item or task. Experience has shown that the reliability of a test is nearly always increased by increasing the length or content of the test. This result is due to the simple fact that, with a large sampling of behavior responses of the person tested, a chance success or failure on one or two items will not appreciably affect the total score. But if a whole test consists of only one item or task, the chance element becomes very important in determining the final test score. Thus with only a single test item such as the Wiggly Block, high reliability could hardly be expected.

---

[47] Johnson O'Connor, *Born That Way* (Baltimore, The Williams and Wilkins Co., 1928).

[48] H. H. Remmers and J. W. Schell, "Testing the O'Connor Wiggly Block Test," *Personnel Journal*, 1933, **12**, 155–159.

Minnesota Spatial Relations Test.[49] In this test the subject is presented with a board about one by three feet, with 58 irregular holes, and with cut-out blocks that fit the holes. He is asked to select the block that fits each hole, and his performance is scored on the basis of time. This test most nearly measures the spatial relations factor, since the subject has to compare the configurations of the blocks with the holes in the board.

Purdue Mechanical Performance Test.[50] This test, a recently developed one,[51] has three subtests, as follows:

Transfer Boards *A* and *B* (these provide for transferring various sizes of pegs and various lengths of U-shaped brackets from one side of a board to the other; this involves judgment of sizes and lengths).

Spatial Relations (this subtest includes four items, namely two "Cue" Boards and two Spatial Relations Blocks).

Hub Assemblies (this subtest consists of eight "Hubs" to be assembled). Pictures of these are shown in Fig. 5.6.

Examples of Mechanical Ability Tests in Use. Mechanical ability tests frequently have been useful in the selection of apprentices. A case in point is one in which the Purdue Mechanical Adaptability Test had been given to 246 railroad-trade apprentices. The apprentices were rated by their instructors and trainers. Those rated as "good" were considered to be performing satisfactory work. The relationship between the scores on the test, and the ratings, is shown in Fig. 5.7, which indicates that more of the apprentices with high test scores were "good" than were those with low test scores.

Case[52] reports a correlation of .54 between supervisors' judgments

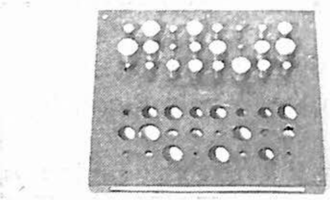
---

[49] Distributed by the Psychological Corporation.

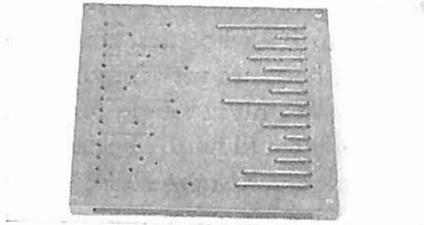
[50] Distributed by the Lafayette Instrument Co.,

[51] Robert L. Brown, *The Development and Validation of a Mechanical Performance Test*, Unpublished Ph.D. Thesis, Purdue University, Jan. 1957.

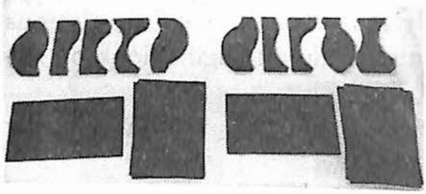
[52] H. W. Case, "Selection of Aircraft Engineering Draftsmen and Designers," *Journal of Applied Psychology*, 1947, 31, 583–588.



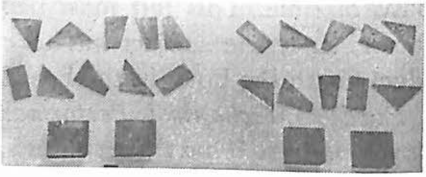
1. Transfer Board A



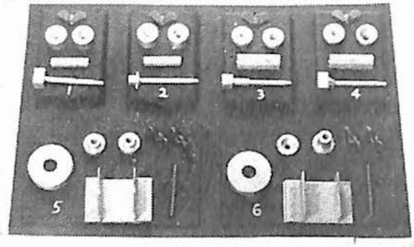
2. Transfer Board B



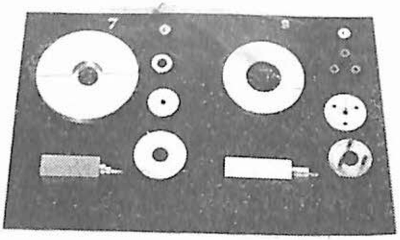
3. Cue Boards A and B



4. Spatial Relations Blocks A and B



5. Hub Assemblies 1–6



6. Hub Assemblies 7 and 8

FIG. 5.6. Purdue Mechanical Performance Test. (Distributed by the Lafayette Instrument Co., Lafayette, Ind.)

of drafting skill and scores made on a test known as the Survey of Space Relations Ability.[53] This test contains a series of geometric figures, each of which is subdivided into parts of different shape. The person taking the test is required in each case to pick out the constituent parts from ten similarly shaped drawings.

In a study of power-sewing-machine operators conducted by Otis,[54]

---

[53] H. W. Case and F. Ruch, *Survey of Space Relations Ability* (California Test Bureau, 1944).

[54] J. L. Otis, "The Prediction of Success in Power Sewing Machine Operating," *Journal of Applied Psychology*, 1938, 22, 350–366.

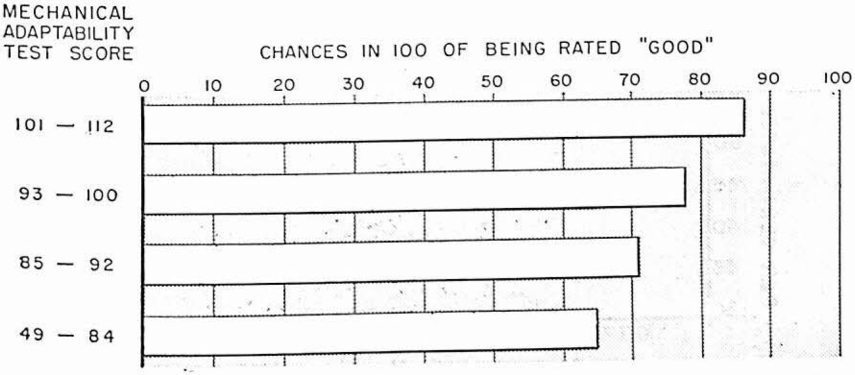


FIG. 5.7. Expectancy chart showing relation between Purdue Mechanical Adaptability Test Scores and rated success of railroad trade apprentices.

it was found that a battery of tests, which included the Minnesota Paper Form Board Test, yielded a multiple correlation of .57 against the quality of work of these operators.

A study using the Purdue Mechanical Adaptability Test[55] on a group of 14 refrigerator-plant mechanics and maintenance men resulted in the scattergram shown in Fig. 5.8. In this study, the rank order correlation between test scores and ratings was .84.

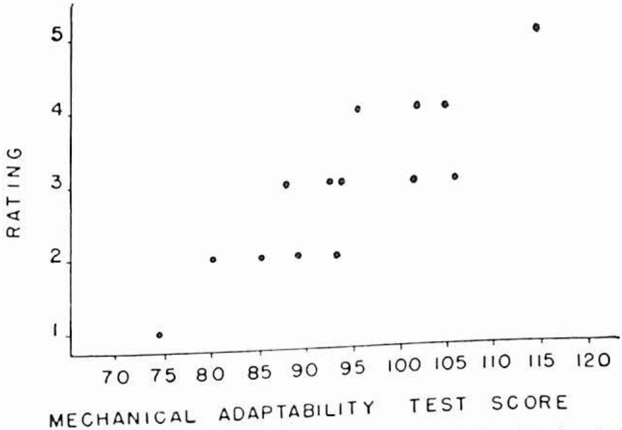


FIG. 5.8. Relation between scores on Purdue Mechanical Adaptability Test and rated ability of 14 refrigerator plant mechanics and maintenance men.

---

[55] Lawshe, Semanek, and Tiffin, *op. cit.*

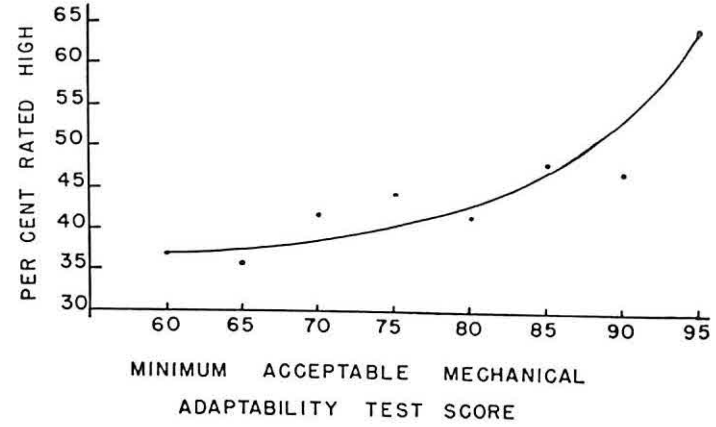


FIG. 5.9. Percentage of automatic screw machine operators who are rated "high" when successively higher test scores on the Purdue Mechanical Adaptability Test are used.

In another study using the Purdue Mechanical Adaptability Test, 46 machine operators from a screw manufacturing company were rated by supervision for their proficiency on the job. A chart showing the percentage of superior operators receiving superior ratings among groups scoring successively higher on the Purdue Mechanical Adaptability Test is shown in Fig. 5.9. This chart indicates that of men scoring 60 or more (a group that included all the operators), only 37 per cent were rated superior, while among operators scoring 95 or above, 65 per cent were rated superior operators.

McMurry and Johnson[56] report an investigation on the selection of welders in which they found a correlation of .66 between job ability of the men and their scores on the Bennett Test of Mechanical Comprehension. Shuman has reported two studies[57] in which significant correlations were found between ability of both employees and supervisors in an airplane-engine-and-propeller plant and their scores on the Bennett Test and Minnesota Paper Form Board. In scoring both operators and supervisors in these studies, Otis Mental Ability Test scores were also found significantly related to job success.

---

[56] R. N. McMurry and D. L. Johnson, "Development of Instruments for Selecting and Placing Factory Employees," *Advanced Management*, 1945, **20**, 113–120.

[57] J. T. Shuman, "The Value of Aptitude Tests for Factory Workers in the Aircraft Engine and Propeller Industries," *Journal of Applied Psychology*, 1945, **29**, 156–160, and J. T. Shuman, "The Value of Aptitude Tests for Supervisory Workers in the Aircraft Engine and Propeller Industries," *Journal of Applied Psychology*, 1945, **29**, 185–190.

Tests of mechanical ability or comprehension are not limited in their application to the machine operating and maintenance jobs of modern industry. A number of investigations have shown that such tests are ideally suited to be a part of an aptitude test battery for the prediction of success in several types of technical training. For example, Laycock and Hutcheon[58] have reported that a test battery consisting of high-school grades, the American Council on Education Psychological Examination, the Cox Mechanical Aptitude Test, the Minnesota Paper Form Board Test, and a questionnaire covering interest in physical science gave a multiple correlation of .66 with later achievement in engineering courses.

Owens[59] describes a new form of the Bennett Test of Mechanical Comprehension (Form CC) that is more difficult than the previous forms. This test, when used with engineering students, "showed a satisfactory correlation with relevant course grades, was an excellent predictor of academic mortality, predicted better at high achievement than at low, made a significant independent contribution beyond that made by the American Council on Education Psychological Examination and high-school average, and gave scores which de-

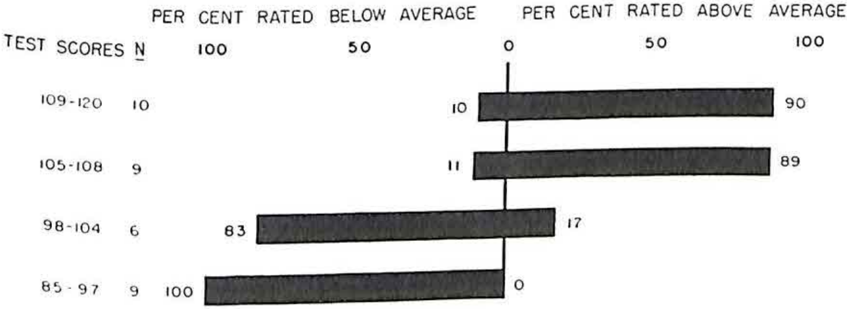


FIG. 5.10. **Relation between scores on the Purdue Mechanical Adaptability Test and ratings of 34 airplane engine mechanic trainees.**

---

[58] S. R. Laycock and N. B. Hutcheon, "A Preliminary Investigation into the Problem of Measuring Engineering Aptitude," *Journal of Educational Psychology*, 1939, 25, 280–289.

[59] W. A. Owens, Jr., "A Difficult New Test of Mechanical Comprehension," *Journal of Applied Psychology*, 1950, 34, 77-81. Reproduced by permission of the American Psychological Association.

pended on training or experience only to the extent of 2 or 3 score points."

The results of an unpublished study on the relationship between scores made on the Purdue Mechanical Adaptability Test and rated success of 34 airplane engine mechanic trainees are shown in Fig. 5.10. The paired comparison rating method (see p. 213) was used in obtaining the ratings by instructors. The percentage of high-rated trainees among men scoring high on the test is very much higher than the percentage of such men among those scoring low on the test.

From such investigations as these we may safely conclude that mechanical aptitude or comprehension tests furnish a real aid to the employment manager in the selection of employees for a wide variety of industrial jobs.

PSYCHOMOTOR TESTS

Psychomotor tests are those which measure muscular abilities, or combinations of sensory and muscular abilities. The term *psychomotor* covers the range of abilities that includes what we commonly call dexterity, manipulative ability, motor ability, and eye-hand coordination, as well as other aspects of muscular performance.

The Nature of Psychomotor Abilities. There is quite a bit of support for saying that there are various *kinds* of psychomotor abilities. This has been supported by the results of studies such as those by Dudek and Seashore,[60] Friedman and Ivens,[61] Harrell,[62] and Seashore, Buxton, and McCollom.[63] In the study by Friedman and Ivens,[64] for example, three different psychomotor factors were identified, as follows:

1. Psychomotor Speed (the ability to carry out quickly and accurately a series of movements that require eye-hand coordination).

---

[60] F. J. Dudek and R. H. Seashore, "Factorial Analysis of Arm-Hand Precision Tests," *American Psychologist*, 1948, 3, 252.

[61] Friedman and Ivens, *op. cit.*

[62] W. Harrell, *op. cit.*

[63] R. H. Seashore, C. E. Buxton, and I. N. McCollom, "Multiple-Factorial Analysis of Fine Motor Skill," *American Journal of Psychology*, 1940, 53, 251–259.

[64] Friedman and Ivens, *op. cit.*

2. Psychomotor Coordination (the ability to coordinate the movements of the large muscles of the body).
3. Finger Dexterity (the ability to manipulate objects rapidly with the fingers).

It should be added that the results of the studies mentioned above, as well as others, are not in complete agreement, so it is not yet possible to present a complete and final listing of the different kinds of psychomotor abilities.

There is also some question as to the extent to which the various psychomotor abilities are *completely* independent, as opposed to being related to each other. In discussing this, however, Vernon[65] points out that in several studies with psychomotor tests, it has been found that such tests typically have some relationships with each other, although the average inter-correlations among groups of them are quite low, illustrations being .25, .27, .30, and .13. Thus, although such tests typically are correlated with each other to some degree, the correlations are so low that psychomotor abilities must be considered as being predominantly specific.

The reader might then raise this question: What are the practical implications of this in using psychomotor tests? The reply would be to the effect that it is *particularly* important in using psychomotor tests to be sure that the test or tests used in a given situation be thoroughly validated for the job in question before they are used. Performance on any one test *cannot* be assumed to give a measure of general psychomotor ability. It is necessary to use specific tests that are *known* to be related to job performance.

Aside from being relatively independent among themselves, psychomotor abilities also are relatively independent of other abilities and characteristics. Let us consider mental abilities, for example. Although the correlations between psychomotor and mental ability tests are somewhat variable (depending on the specific tests and samples of people), the correlations tend to be low, or zero. For example, in an investigation by the Western Electric Company,[66] the correlation between the O'Connor Finger Dexterity (described below) and the Otis Mental Ability Test was found to be .07. The correlation between the Hayes Pegboard (another dexterity test described below) and the Otis Test was found to be zero. Each of these correlations was based on test scores of 749 employees. These results are typical of the results of many similar investigations. Although the correlations generally tend to be positive rather than zero, they typically are so low as to be of little practical value.

---

[65] P. E. Vernon, "The Structure of Human Abilities," (London, Methuin and Co., Ltd., and New York, John Wiley and Long, 1950).

[66] *Analysis of 1935–37 Experience in Selecting New Men for Shop Occupations.* Privately printed Monograph (Chicago, Western Electric Company, Hawthorne Plant, 1939).

Somewhat the same can be said about the relationship between psychomotor tests and mechanical ability tests. In the study by Friedman and Ivens,[67] for example, the following correlations were reported for the specified psychomotor tests with a test of Mechanical Principles: Circle Dotting, .02; Two-Hand Drawing, .03; Two-Hand Coordination, .29; Rotary Pursuit, .19; and Finger Dexterity, .11. These very low correlations justify the conclusion that, for practical purposes, psychomotor abilities are relatively independent of mechanical abilities.

Neither is there any appreciable relation between psychomotor abilities and anthropometric measurements, such as length of fingers or other body measurements. This conclusion was reached by Griffitts[68] and has been corroborated by other investigators.

Typical Psychomotor Tests. *O'Connor Finger Dexterity Test.*[69] This test, illustrated in Fig. 5.11, is a widely used manipulative test. The equipment consists of a metal plate of 100 holes, and 310 small cylindrical pins. After a brief instruction period, the subject is asked to fill the board—three pins to a hole—as quickly as possible. The frequency distributions shown in Fig. 2.3, p. 18, were obtained by means of the O'Connor Finger Dexterity Test. These distributions show that the time required to fill the board varies from around five to fifteen minutes, depending upon the amount of finger dexterity possessed by the applicant.

Hayes Pegboard.[70] This test, illustrated in Fig. 5.12, also measures finger dexterity but involves, in addition, an element of hand and arm co-ordination. The board contains two rows of 26 holes each.

---

[67] Friedman and Ivens, *op. cit.*

[68] C. W. Griffitts, "The Relation Between Anthropometric Measures and Manual Dexterity," *Journal of Applied Psychology*, 1936, **20**, 227–235.

[69] Distributed by the Lafayette Instrument Co.

[70] Elinor G. Hayes, "Selecting Women for Shop Work," *Personnel Journal*, 1932, **11**, 69–85.

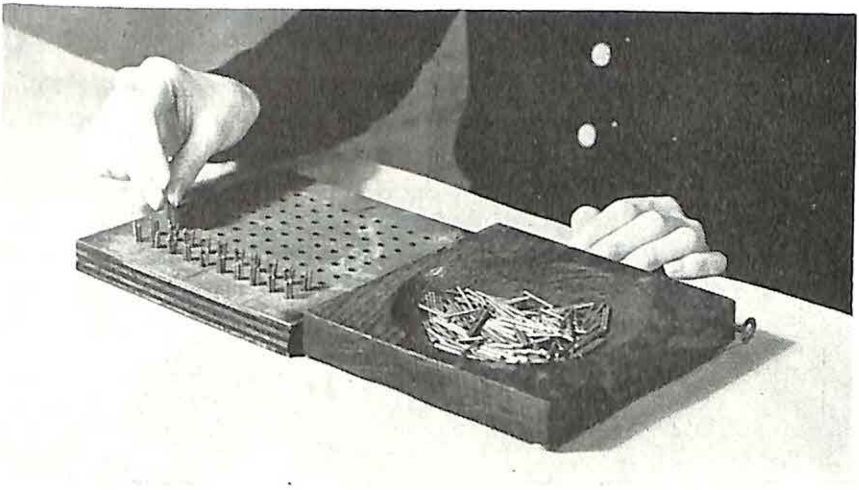


FIG. 5.11. O'Connor Finger Dexterity Test.

The administration calls for a brief practice period, after which the person being tested is given nine trials of 30 seconds each, three of these trials being with the right hand, three with the left, and three with both hands. The score is the total number of pins placed during the nine trials.

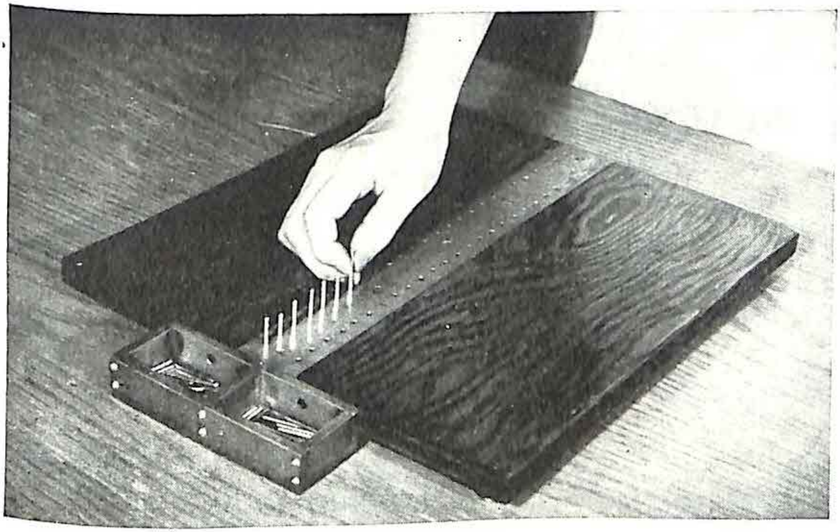


FIG. 5.12. The Hayes Pegboard.

Purdue Pegboard.[71] This test, which is illustrated in Fig. 5.13, measures separately two basic aspects of manipulative dexterity. The assembly of a series of pin-collar-washer arrangements measures fine finger dexterity of the type measured by the O'Connor Finger Dexterity Test. The placing of pins into a series of holes measures manual dexterity of the type involved in the Hayes Pegboard. These two measurements may be obtained with the same board and require only two and one-half minutes of testing time. The Purdue Pegboard gives separate measurements for right hand, left hand, and both hands.

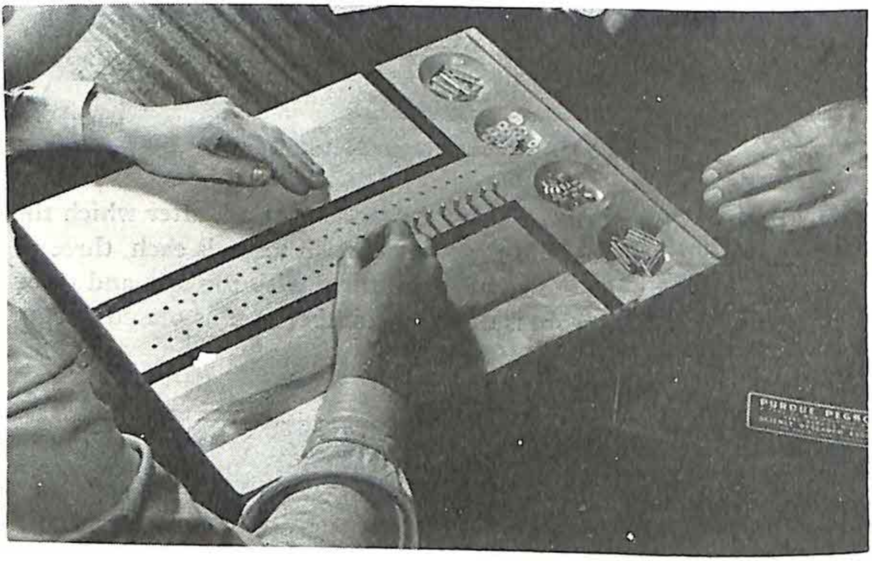


FIG. 5.13. Purdue Pegboard.

A further advantage of the Purdue Pegboard is that it can be given to ten or more persons simultaneously. All that is required for each group testing is a test board for each person and the careful attention of the test administrator.[72]

---

[71] Distributed by Science Research Associates, Chicago, Illinois. For reference, see Table 5.2 on p. 142. For a more detailed discussion of this test see J. Tiffin and E. J. Asher, "The Purdue Pegboard: Norms and Studies of Reliability and Validity," *Journal of Applied Psychology*, 1948, 32, 234–247.

[72] In one industrial employment office this test has been successfully given by a single examiner at the rate of 50 applicants per hour. Ten test boards were used simultaneously in this office.

Minnesota Rate of Manipulation Test.[73] This test consists of 58 round blocks 1.5 inches in diameter, and a board 10 by 35 inches with holes slightly larger than the blocks. The blocks are slightly thicker than the board. Two tests may be given with this equipment—namely, the placing test and the turning test. The placing test requires that all the blocks be placed in the holes in the board. The turning test requires that all the blocks be turned over. In addition to the early descriptions of this test, later work on norms and on a modified method of administration have been published, respectively, by Cook and Barre[74] and by Jurgensen.[75]

Stromberg Dexterity Test.[76] This test consists of equipment somewhat like that of the Minnesota Rate of Manipulation Test. The blocks are colored red, blue, or yellow, and instead of repeating the same simple motion, each block has to be moved differentially from every other block, depending on the color. Thus, the test requires both discrimination and placement. It takes five to ten minutes to administer the test.

MacQuarrie Test for Mechanical Ability.[77] This test, in spite of its name, measures, in part, various psychomotor abilities. Unlike most psychomotor tests, this one is a paper-and-pencil test, and it lends itself to group testing. There are seven parts of the test: speed and accuracy of tracing, tapping, dotting, copying, letter location, block identification, and visual pursuit. The first three of these deal primarily with psychomotor abilities. Illustrations of three parts are given in Fig. 5.14.

Other dexterity tests. A dexterity test that worked well in placing operators on the job of looping in the hosiery industry makes use of a grooved pegboard. This test calls for a board containing five rows of five holes each. The holes are .125 inch in diameter and are spaced

---

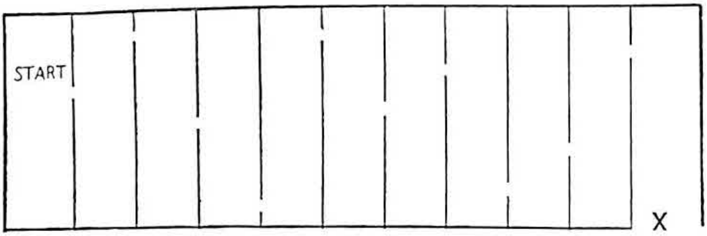
[73] Distributed by the C. H. Stoelting Co., 424 N. Homan Ave., Chicago, 24, Ill. Described by H. J. Green, I. R. Berman, D. G. Paterson, and M. R. Trabue, "A Manual of Selected Occupational Tests," *Bulletin of the Employment Stabilization Research Institute*, 1933, 2, 59–85.

[74] D. W. Cook and M. F. Barre, "The Effect of Specialized Industrial Norms on the Minnesota Rate of Manipulation Test as a Selective Instrument in Employment Procedure," *Journal of Applied Psychology*, 1942, 26, 785–792.

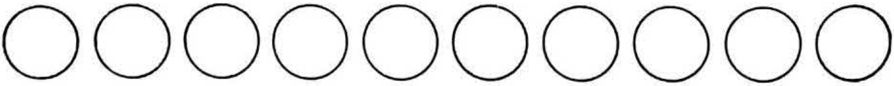
[75] C. E. Jurgensen, "Extension of the Minnesota Rate of Manipulation Test," *Journal of Applied Psychology*, 1943, 27, 164–169.

[76] Distributed by the Psychological Corporation.

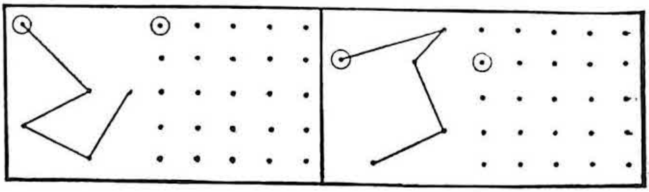
[77] T. W. MacQuarrie, *MacQuarrie Test for Mechanical Ability* (California Test Bureau, 1925).



Tracing: Applicant draws a line from "start" to X without crossing a vertical line.



Dotting: Applicant places three dots in each circle as quickly as possible.



Copying: Applicant reproduces the drawings by connecting the appropriate dots in the area to the right of each drawing.

FIG. 5.14. Several parts of the MacQuarrie Test for Mechanical Ability. This test is essentially a group test of muscular co-ordination and control.

½ inch apart. On one side of each hole a groove has been cut. The location of these grooves is varied randomly from one hole to another. Each of the pins to go in the holes has a key along one edge so that the pin may be inserted in the hole only when the key and groove are located correctly with respect to each other. This test is referred to as the Purdue Grooved Pegboard.[78] The test is scored by determining the average time required to fill the board for the third and fourth trials of four that are allowed.

Many other tests of dexterity, muscular co-ordination, or manipulative ability are available. Some of these are specifically adapted to certain jobs and are of little value for other types of placement.

---

[78] The source of this test is given in Table 5.2 on p. 142.

Others have been found suitable as placement tests for several different kinds of jobs. Before accepting any one of these tests as suitable for any given job, one should be sure the test actually "comes through" in an evaluation of the type described on pp. 81–90. The names and sources of some of the more commonly used tests of this type are summarized in Table 5.2.

Examples of Psychomotor Tests in Use. Candee and Blum[79] have reported a study of the relation between finger and tweezer dexterity and the efficiency of workers in a watch factory. They found a statistically significant difference between average scores on the O'Connor Finger Dexterity Test of the superior and mediocre workers. This relationship is represented by a correlation of .26 between foremen's ratings and test scores. The fact that the obtained correlation was no higher was due, in all probability, to the unreliability of the foremen's ratings. Even with a validity coefficient of .26, however, the test will have real value as a placement device if a sufficiently small selection ratio is utilized. A tweezer dexterity test (a test in which a pair of tweezers instead of the fingers is used in placing the pins), on the other hand, showed practically no relationship with foremen's ratings. One might conclude, *a priori*, that since much of the work in a watch factory is done with a pair of tweezers, a test involving tweezer dexterity would be a more desirable test for this type of work than one involving finger dexterity. The experimental facts, however, support the opposite conclusion. This case will serve to illustrate again the fact that tests should not be accepted merely because of logical considerations in the absence of experiments that clearly show whether they will or will not work for the job in question.

A study of the O'Connor Finger Dexterity Test in selecting electrical-fixture assemblers has been reported by Tiffin and Greenly.[80] After experience on the job had been ruled out by partial correlation, the test scores were found to correlate with productivity, as indicated by earnings, to the extent of .22, and to correlate with general efficiency as indicated by merit ratings to the extent of .33. Although these validity coefficients are not so high as might be desired, the

---

[79] B. Candee and M. Blum, "Report of a Study Done in a Watch Factory," *Journal of Applied Psychology*, 1937, **21**, 572–582.

[80] J. Tiffin and R. J. Greenly, "Employee Selection Tests for Electrical Fixture Assemblers and Radio Assemblers," *Journal of Applied Psychology*, 1939, **23**, 240–263.

TABLE 5.2

Dexterity and Muscular Co-ordination Tests

Name of test	*What is measured*	*Source*	*Reference*
Minnesota Rate of Manipulation	Finger and Hand Dexterity	Psychological Corporation, 304 East 45th St., New York City	H. J. Green, R. I. Berman, D. G. Paterson and M. R. Trabue, "A Manual of Selected Occupational Tests," *Bulletin of the Employment Stabilization Research Institute*, 1933, **2**, 59–85.
O'Connor Finger Dexterity	Finger Dexterity (fine movements)	Lafayette Instrument Co., 52 By-Pass, Lafayette, Indiana	Mildred Hines and Johnson O'Connor, "A Measure of Finger Dexterity," *Personnel Journal*, 1926, **4**, 379–382.
Purdue Pegboard	Finger Dexterity (fine movements) and Finger Dexterity (gross movements)	Science Research Associates, 57 W. Grand Ave., Chicago, Ill.	Manual of Instructions for the Purdue Pegboard, Science Research Associates, 57 West Grand Ave., Chicago, Ill.
O'Connor Tweezer Dexterity	Dexterity in Using Tweezers	Lafayette Instrument Co., 52 By-Pass, Lafayette, Indiana	Johnson O'Connor, *Born That Way* (Baltimore, Williams and Wilkins Co., 1928).
Stanford Motor Skills Unit	A Variety of Motor Skills	C. H. Stoelting & Co., 424 N. Homan Ave., Chicago, Ill.	R. H. Seashore, "Stanford Motor Skills Unit," *Psychol. Monog.*, 1928, **39**, No. 2, 51–66.
Purdue Grooved Pegboard	Specialized Finger Dexterity	Lafayette Instrument Co., 52 By-Pass, Lafayette, Indiana	p. 140 of this book.

statistical chances are ninety-eight out of one hundred that even the lower one represents a real relationship between the test scores and production. If the selection ratio is kept small, say in the neighborhood of .10, one may reasonably expect the test to be of definite value in placing employees on this job.

An experiment showing the value of dexterity tests in selecting operators of punch presses (and similar machines) has been reported by Hayes.[81] Hayes used two psychomotor tests—the Hayes Pegboard and the O'Connor Finger Dexterity Test—with a scoring system that gave each test its optimum weight in deriving a composite score for each operator. On the basis of output during the first eight weeks, the operators were divided into the following four criterion groups: quick learners, fair learners, slow learners, and those who did not complete the training. The relationship between the composite scores on the two tests, and these criterion groups, is shown in Fig. 5.15. This shows quite clearly that the two tests would be of material benefit in identifying "quick" learners for this type of work. Somewhat comparable results were obtained with three other jobs.

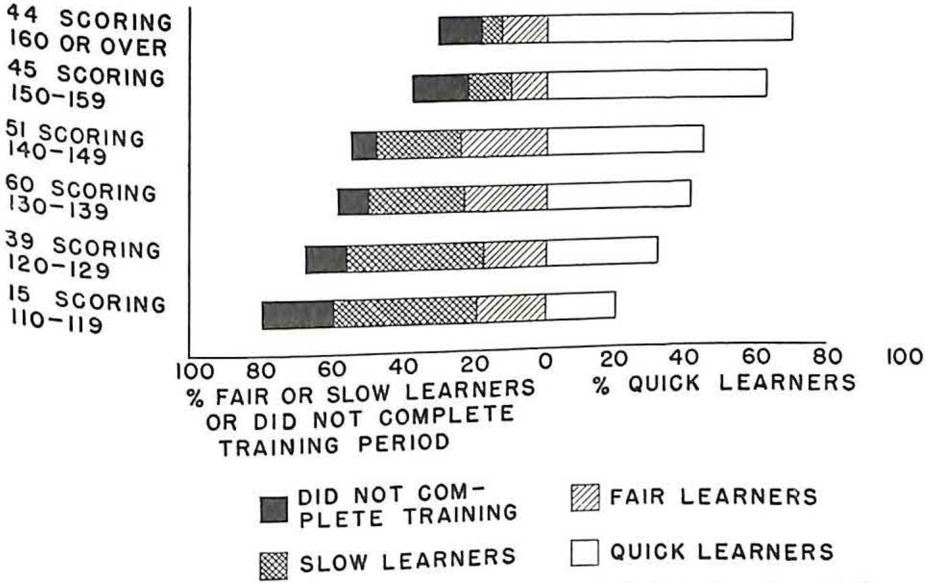
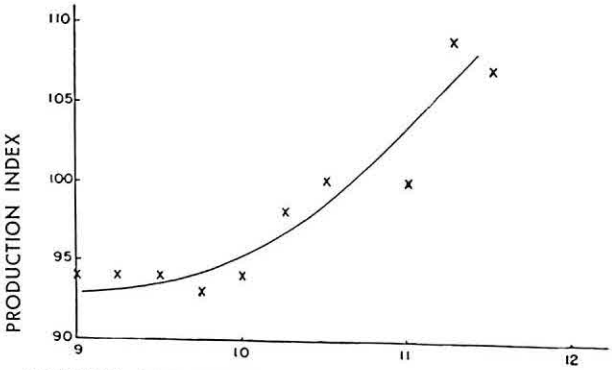


FIG. 5.15. Relation between composite score on two dexterity tests and speed of learning for a group of 254 operators of punch presses and similar machines.

---

[81] Hayes, *op. cit.*



MINIMUM ACCEPTABLE PURDUE PEGBOARD ASSEMBLY SCORE

FIG. 5.16. **Job performance of "quillers" in relation to successively higher scores on the assembly part of the Purdue Pegboard. (From data supplied courtesy the Burlington Mills Corporation, Burlington, N.C.)**

An unpublished study on the use of the Purdue Pegboard with employees on an operation known as "quilling" in a textile mill resulted in a curve plotted in Fig. 5.16. Production indices of 28 operators of approximately equal experience were obtained. In Fig. 5.16 the production indices are plotted against minimum assembly scores on the Purdue Pegboard. As the curve shows, among the operators scoring 9 or above (a category including all the operators) the average production index was 94; with successively higher minimum scores, the average production index was also successively higher.

Trade apprentices need much more than capacity for developing the manipulative aspects of their trade. They need, first of all, capacity for learning the fund of knowledge that is required of an expert tradesman. This capacity is better measured by tests of mental ability and mechanical comprehension than by tests of dexterity and muscular co-ordination. Yet, in spite of the fundamental importance of these mental capacities, the element of dexterity is also of some importance in achieving success as a trade apprentice. Allen,[82] for example, reports that a battery of seven tests covering intelligence, mechanical aptitude, mechanical ability, and dexterity was found of value in selecting boys who are able to make satisfactory progress in learning the various skilled trades. Although three of the four elements measured in this study deal with the mental aspects of the trainees, it is significant to note that dexterity also was of some importance.

---

[82] E. P. Allen, "The Selection of Engineering Apprentices," *Journal of the National Institute of Industrial Psychology*, 1931, 5, 379–384.

VISION TESTS

Practically every industrial job requires some degree of vision, and many jobs require a high degree of skill in some particular visual function. The inspector of small parts for "appearance" must have keen vision at close distances. The operator of certain knitting and other textile machines must not only have keen vision at close distances but also must be able to maintain such vision for long periods with only occasional interruption. The truck driver, crane operator, and signal man must have keen vision for greater distances and good perception of space relationships. Color discrimination is of importance to an employee wiring a radio (because he must discriminate between wires of different color), to the operator of a color-printing press, to the pipefitter tracing a color code through the plumbing system of a plant, and to the operator of mobile equipment who must depend upon colored signals to determine whether roads are open or pathways clear. Various measurable visual characteristics have been found to be related to successful performance on certain jobs—even on jobs in which these visual factors could not have been inferred from ordinary job analysis procedures. The pattern of visual requirements differs markedly for different jobs, and these requirements cannot be discovered except by a study of the people on the job, as explained in Chapter 4.

Management in production industries has long recognized the importance of good vision in employees.[83] Some form of vision test, administered either in the employment office or as a part of the medical examination, is perhaps more common in industry than is any other form of employee test. Testing visual skills has always been a legitimate psychological function along with testing carried on in other areas of skill or aptitude.

Visual skills are not innate and permanent; they are to a large degree acquired and can be, and often are, modified. They change universally with age. Professional eye care, based on a clinical or diagnostic examination, can often give to an employee the visual skills he needs for his job or restore the skills he has lost with increasing age.

---

[83] Extensive recent publications in this field are listed in bibliographies on industrial vision:

H. S. Kuhn, "Articles Bearing on Industrial Eye Problems," *Transactions of the American Academy of Ophthalmology and Otolaryngology*, 1946, **50**, 175–178.

N. F. Stump, "A New Concept of Ophthalmic Service," *Bausch & Lomb Magazine*, 1946, **22**, No. 1, 4–20.

Although our primary interest in vision tests is for personnel selection and placement purposes, it should be recognized that such tests are also used for other purposes. One of these is in connection with the maintenance of employees' visual skills. If, by the use of a vision test, it is found that an employee's vision is deficient for his job, he may be referred to a physician or optometrist for professional eye care. Another use is in connection with workmen's compensation. For this purpose it is necessary to have some measure of the individual's visual skills. Thus, although we will be discussing the use of vision tests for aptitude testing purposes, their other uses should also be kept in mind.

The Nature of Visual Skills. Extensive research on the visual skills of people on various jobs in industry and business has indicated certain aspects of visual performance as being most important for classifying and placing employees according to differences in visual characteristics. These aspects of visual performance are described below, along with some observations regarding tests for measuring such skills.

Keenness of vision (visual acuity) at appropriate distances (usually tested at 20 feet and 13 or 16 inches). This visual function is the ability to discriminate black and white detail, measured in terms of the minimum separable areas that can be distinguished. For industrial placement such a test should be equally valid for illiterate and literate subjects and should avoid the complications introduced by a factor involving discrimination and recognition of different shapes such as letters. The scale should make it possible to classify acuity scores adequately throughout the entire range of performance. Since acuity is modified by brightness, glare, and other external conditions, the acuity test must be given under controlled and standard conditions.

Discrimination of differences in distance (depth perception, or stereopsis). This function is an important phase of correct perception of spatial relationships. Of several cues for judging relative distances of objects, the most important for persons with normal vision, and the one that can be controlled and measured most reliably, depends on the slight difference in the position of the two eyes. The two eyes perform a geometric triangulation upon a distant object, and the distance of that object is perceived through an integration of the minute differences in appearance of the object to the two eyes. Other cues for perceiving distance in the third dimension may augment but cannot adequately substitute for this cue from two-eye functioning. Stereopsis is measurable quantitatively like most other human functions and should be measured for employee placement.

Discrimination of differences in color. Accurate color discrimination is important for efficient performance on many industrial jobs. The particular colors to be differentiated should include as many combinations as there are factors in color sensitivity. The most important combinations are those that represent common colors in signal lights. In order to read such signals correctly, it is important not only that an employee should be able to differentiate between them but also that he should be able to identify and interpret correctly the meaning of each color. Any test of color vision can and should be scaled for quantitative measurement and classification of employees.

Postural characteristics of the eyes (phorias) at appropriate distances (usually 20 feet and 13 or 16 inches). Under normal seeing conditions the two eyes must move in relation to each other so that both converge symmetrically upon the object. It is this convergence that gives us a clear, single image of the object when it is viewed binocularly. Under certain testing conditions, which eliminate the necessity for such convergence of the eyes on a single point, the eyes assume a posture that may converge or diverge from that required in normal seeing at the test distance. Such postures (called "phorias" in clinical terminology) are measured in terms of angular deviation from the posture normally required for that distance. The deviation may be lateral or vertical and is measured separately in each direction. The explanation of this phenomenon has not been determined finally; nevertheless, such characteristics should be measured because they are related to performance on many industrial jobs. Such measurement must be done with adequate control and standardization of the several factors that may modify the measurement, such as the distance and focus requirements of the test.

These are not, of course, the only visual functions that are of importance in all industries, but they comprise a reasonable minimum for a program of testing applicants and employees. Tests of these functions have been adapted satisfactorily for simple and rapid use in industry, and they are the ones most widely recommended by leaders in industrial eye care[84] for use in industry.

Examples of Vision Tests. Probably the most commonly used vision test is the Snellen letter chart that is a familiar sight on the walls of physicians' offices. The test consists of several rows of block letters of decreasing size, usually placed at a distance of 20 feet from the subject. A typical Snellen chart is shown in Fig. 5.17. The test is administered by determining, separately for each eye, the smallest letters that the subject can read. The larger the letters at 20 feet that are the smallest readable, the poorer the visual acuity. The Snellen notation of acuity

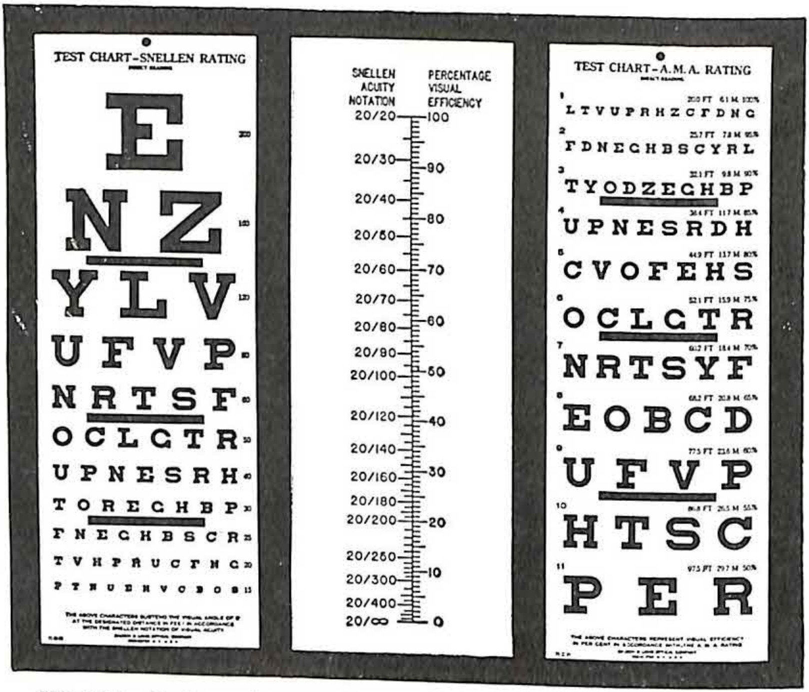


FIG. 5.17. Snellen and "A.M.A." test charts with scale for converting Snellen scores to per cent of visual efficiency.

---

[84] H. S. Kuhn, *A New Concept of Visual Performance in Industry*, Publication No. 340 (National Society for the Prevention of Blindness, 1940).

A. C. Snell, "The Field of Industrial Ophthalmology," American Medical Association, Section of Industrial Health (Chicago, Jan. 13, 1942).

C. F. Shepard, "Visual Skills," *Optometric Weekly*, 1944, 34, 1465–1466.

scores is in the form of a fraction—the smaller the fraction, the poorer the vision. In this fraction the numerator is constant and represents the distance of the test. Thus, visual acuity scored 20/20 is standard. A score of 20/40 means that the subject can read at 20 feet only a letter twice as large as standard, a letter that the "standard eye" can read at 40 feet.

The Snellen acuity designations are not intended to represent fractions of useful vision. The employee who scores 20/40 at a distance of 20 feet is not necessarily handicapped 50 per cent in opportunity to earn a living. In order to simplify interpretation of visual acuity scores and to set up an equitable scale for awarding compensation in proportion to actual incapacity due to eye injury, the American Medical Association has adopted and recommended for use in industry a percentage system of acuity notation, with the distinguishing title of "Visual Efficiency."[85] Where the Snellen test measures acuity from the angles subtended by the letters, the American Medical Association notation interprets these angles in terms of percentage of visual efficiency. The difference between this percentage and 100 per cent is the percentage "loss of vision." A conversion scale for translating acuity scores into the percentage notation is shown in Fig. 5.17. Also shown is the first half of a letter chart that measures visual angles in steps directly equivalent to intervals of 5 per cent on the American Medical Association scale.

This American Medical Association percentage notation is ordinarily not carried above 100 per cent or below 20 per cent, and for compensation purposes in industry no extension of this range is necessary. Acuity above standard is usually considered as only 100 per cent in computing compensation awards; and in the absence of records, 100 per cent acuity is assumed prior to the time of an eye injury. Twenty per cent acuity or less is frequently considered as complete lack of acuity or industrial blindness, and is therefore not measured or classified more precisely. These practices vary in different states, and the industrial relations officer should be familiar with the legal practices in his state regarding compensation for eye injury in industry.[86]

---

[85] Albert C. Snell and Scott Sterling. "The Percentage Evaluation of Macular Vision," *Archives of Ophthalmology*, 1925, **54**, 443–461.

[86] See A. C. Snell, *Medicolegal Ophthalmology* (St. Louis, C. V. Mosby Company, 1940).

In the upper range of acuity scores (standard or above) adequate segregation of differences in acuity is not possible with these letter charts, since the common Snellen charts provide only two or three test levels above standard and the percentage chart provides none. Other limitations of the Snellen letter test have been pointed out,[87] and nonliterate test charts have been developed. These charts eliminate letters and require the subject to identify a spatial pattern, such as a broken ring that may have its open area at the top, bottom, right, or left. In spite of the limitations of both the Snellen and American Medical Association charts, these tests have furnished industry with a convenient and acceptable basis for segregating and classifying cases of substandard acuity in questions of compensation for eye injury. Equivalent information may be obtained, of course, from more elaborate vision tests used in the plant for other purposes.

Because such tests as the Snellen and the American Medical Association do not adequately serve selection and placement purposes, various other vision-testing devices have been developed. Such devices measure various visual skills in addition to visual acuity, and also provide for the testing of people at the optical equivalent of both near and far distances. Three such instruments are the Ortho-Rater, the Sight-Screener, and the Telebinocular.[88]

The Bausch & Lomb Visual Classification and Placement Tests for Industry[89] are the first battery of vision tests to be constructed on the basis of specifications derived from extensive investigations among industrial employees in industrial situations. These tests cover the visual functions described previously and, for maximum speed and convenience in testing, are incorporated in a single instrument, shown in Fig. 5.18. This instrument, called the Ortho-Rater, is a precision stereoscope of relatively long focal length that permits adequate and separate control of test stimuli for each eye. Tests are given at optical equivalents of 26 feet and 13 inches. Stereoscopic methods of vision testing have been used since the late nineteenth

---

[87] C. E. Ferree and G. Rand, "A New Method of Rating Visual Acuity," *Journal of General Psychology*, 1941, **25**, 143–176.

[88] The Ortho-Rater is manufactured by Bausch & Lomb Optical Company, Rochester, New York; the Sight-Screener by American Optical Company, Southbridge, Massachusetts; the Telebinocular by Keystone View Company, Meadville, Pennsylvania.

[89] The Bausch & Lomb Optical Company, Rochester, New York.

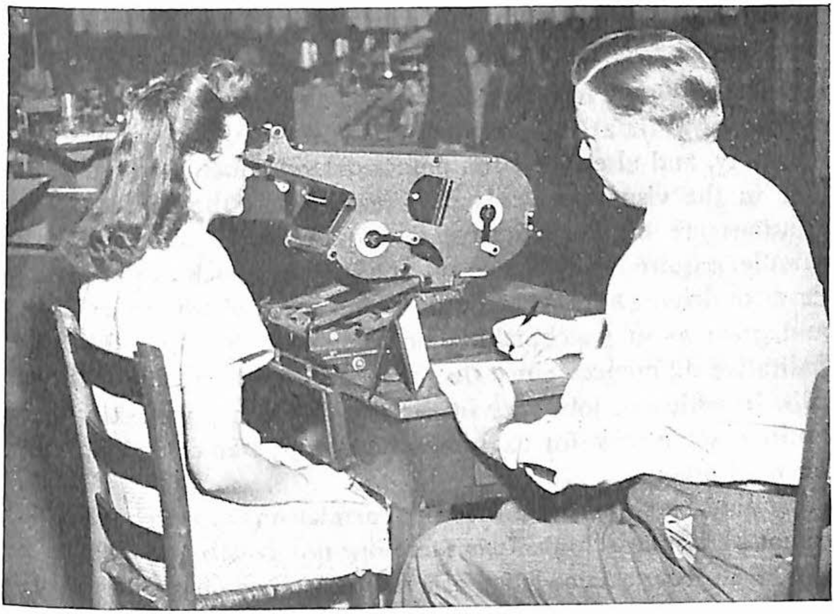


FIG. 5.18. The Ortho-Rater in use in a hosiery mill. (Courtesy Haynes Hosiery Mills Company, Winston-Salem, N.C.)

century and were early described by Wells.[90] The statistical data upon which these tests are based have been published separately in technical and professional journals.[91]

The Ortho-Rater includes the following twelve tests: far-distance tests of vertical phoria, lateral phoria, acuity both eyes, acuity left eye, acuity right eye, depth perception, and color discrimination;

---

[90] David Wells, *The Stereoscope in Ophthalmology* (2nd ed.; Globe Optical Company, 1918).

[91] W. J. Giese, "The Interrelationship of Visual Acuity at Different Distances," *Journal of Applied Psychology,* 1946, **30,** 91–106.

F. W. Jobe, "Instrumentation for the Bausch & Lomb Industrial Vision Service," *Bausch & Lomb Magazine,* 1944, **20,** No. 2, 3–5, 19–21.

Joseph Tiffin and H. S. Kuhn, "Color Discrimination in Industry," *Archives of Ophthalmology,* 1942, **28,** 851–859.

Joseph Tiffin and S. E. Wirt, "Near *vs.* Distance Visual Acuity in Relation to Success on Close Industrial Jobs," *Supplement to Transactions of the American Academy of Ophthalmology and Otolaryngology,* June 1944, 9–16.

S. E. Wirt, "The Validity of Lateral Phoria Measurements in the Ortho-Rater," *Journal of Applied Psychology,* 1943, **27,** 217–232.

near-distance tests of vertical phoria, lateral phoria, acuity both eyes, acuity left eye, and acuity right eye.

Examples of Vision Tests in Use. Extensive research carried out over a period of years at the Occupational Research Center of Purdue University, and elsewhere, has demonstrated conclusively that jobs differ in the visual demands they make upon the worker. These variations are both qualitative and quantitative. Some jobs, for example, require the ability to see at a distance, such as in operating a crane or driving a truck; and others require the ability to see close at hand, such as in watchmaking or fine assembly work. These are qualitative differences, since they are predicated on different visual skills. In addition, jobs vary in the amount of a given skill that is required; some jobs, for example, require greater color discrimination than others.

It should be added here that the correlations among visual skills are not particularly high. It is therefore not possible to predict one visual skill from a knowledge of a person's score on a different type of vision test. Being able to read the chart across the doctor's room gives no assurance of being able to read the fine print on the insurance policy.

In the industrial vision research conducted at the Occupational Research Center, it was found that vision tests could be most effectively used if visual standards for jobs were set up in the form of "visual profiles," each profile specifying the minimum acceptable scores on the various Ortho-Rater tests. If a person had scores above the specified minimums, he "passed" the profile, and if he had one or more scores below any specified minimum, he "failed" the profile. Profiles for individual jobs were developed, using various criteria such as quantity of work, quality of work, ratings, and accidents.

Visual job families. Over a period of years data were collected on vision test scores and criteria of job success for individuals on several thousand jobs. A careful analysis of the relationships between vision test scores and job success as they have been revealed in these many different job situations has shown that there are in industry certain groups of jobs that are similar to each other in terms of visual requirements. Within each group the visual requirements of the jobs are similar, but the requirements of each group vary from those of the next group. Thus we have a series of job groupings each representing a different pattern of visual requirements although the jobs within each group are essentially similar in visual requirements. This has led to the concept of "visual job families." A visual job family is composed of a group of jobs whose visual requirements are similar. The requirements between families, however, differ.

Six such visual job families have been identified. Research has shown that the vast majority of industrial jobs will fall into one or another of these six groups in terms of the visual demands the job makes upon the worker. The visual requirements in terms of Ortho-Rater test scores of each of the six job families are shown in Fig. 5.19. Descriptions of these families are given below:

Visual Job Family I. Clerical and administrative. This standard covers those jobs primarily concerned with paper work. All types of clerical jobs and those administrative occupations that are of the desk-work type are included.

Visual Job Family II. Inspection and close work. This standard covers jobs involved in the inspection of small parts for surface defects. Also involved are jobs of the machine-operating type in which the work is done at close range (such as sewing-machine operator). Assembly jobs involving very small parts (such as watches, radio tubes, and so forth) also fall in this category.

Visual Job Family III. Vehicle operator. This standard covers jobs requiring the operation of moving vehicles (truck driver, crane operator, high-lift operator, and so forth).

Visual Job Family IV. Machine operator. This standard covers those jobs involving the operation of machines in which the operating parts of the machine are within arm's length (such as lathes, drill presses, spinning machines, and so forth).

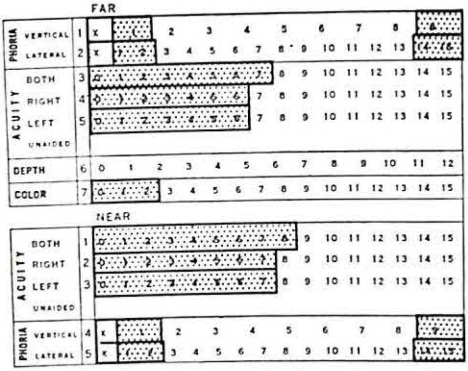
Visual Job Family V. Laborer. This standard involves jobs of the relatively unskilled type (porters, janitors, guards, hand truckers, and so forth).

Visual Job Family VI. Mechanic and skilled tradesman. This standard involves jobs of the mechanical type (such as radio mechanic, Diesel mechanic, machine fixer, and so forth). Also included are the skilled trades (such as carpenter, plumber, millwright, electrician, and so forth).

In order to investigate the relationship between visual skill test score requirements as represented by these six visual standards and success on industrial jobs the method of cross-validation was used.

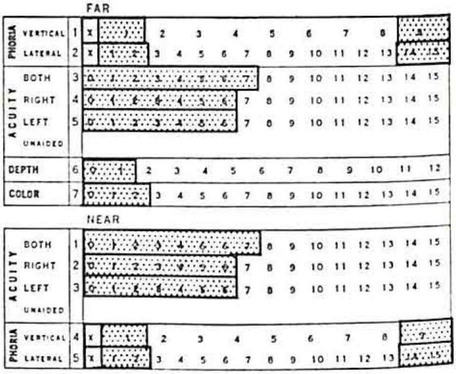
CLERICAL AND ADMINISTRATIVE

VISUAL PERFORMANCE PROFILE



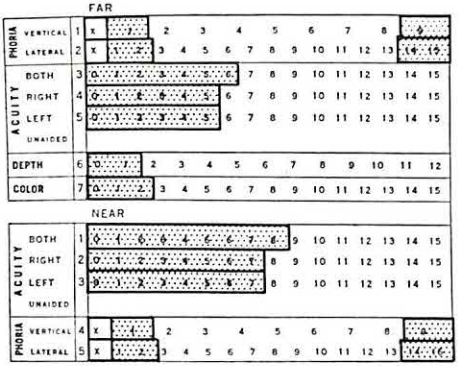
MACHINE OPERATOR

VISUAL PERFORMANCE PROFILE



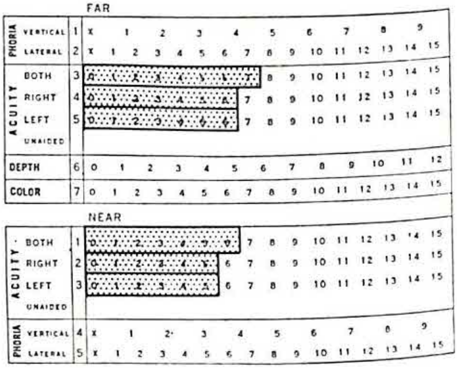
INSPECTION

VISUAL PERFORMANCE PROFILE



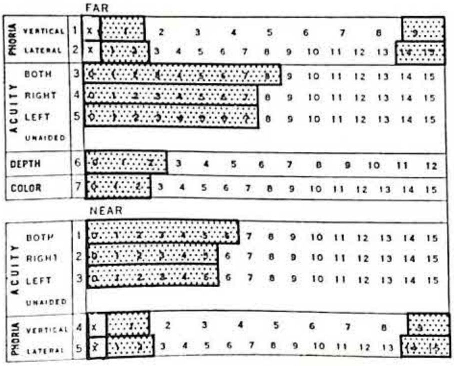
LABORER

VISUAL PERFORMANCE PROFILE



VEHICLE OPERATOR

VISUAL PERFORMANCE PROFILE



MECHANIC

VISUAL PERFORMANCE PROFILE

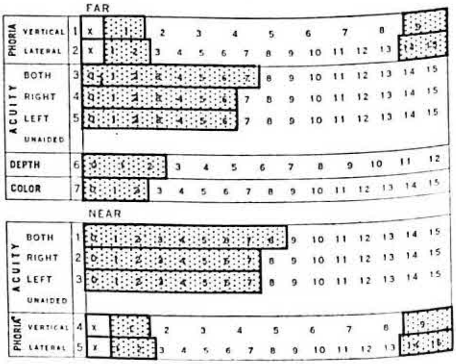


FIG. 5.19. Visual skill test requirements for the most common job groups. An employee "passes" the visual requirements if all of his Ortho-Rater scores fall in unshaded areas. He "fails" if one or more scores fall in a shaded area.

A random sample of 43 individual job studies was selected. Each unit of this sample represented a group of data involving visual test scores and measures of job success. From the job-description materials supplied by the companies, the jobs were assigned to one of the six visual job families. The workers were then divided into those whose visual skills were adequate for the job and those whose visual skills were not adequate for the job, using the visual skills profile appropriate for this job family. For each of these visual groups the percentage of workers who were in the high-criterion group on the basis of job success was computed. Thus each unit in this sample represented an individual study of the relationship between vision and job success in an individual job situation in a specific plant.

Table 5.3 presents the results of these studies grouped by visual job families. This table shows the number of jobs classified into each of the six job families, the number of employees in each of these groups, and the per cent of "high-criterion" employees in each group among those whose vision is adequate for their jobs and the corresponding per cent of "high-criterion" employees among those whose vision is inadequate for their jobs.

The obvious conclusion from the results shown in Table 5.3 is that there is consistently a higher per cent of high-criterion employees among those whose vision is adequate for their jobs than among those whose vision is lacking in this respect. It should be kept in mind that satisfactory vision for one visual job family does not at all guarantee adequate vision for another type of job.

TABLE 5.3

Relationship between Visual Skill Profiles and Success on Industrial Jobs in Each of Six Visual Job Families

Visual Job Family	*No. of Jobs*	*No. of Employees in Job Family Group*	*Per cent of "high-criterion" employees on the jobs among those who have:*	
			Adequate Vision	*Inadequate Vision*
I	5	170	71%	37%
II	9	783	62%	50%
III	5	391	59%	45%
IV	10	497	63%	45%
V	5	234	67%	34%
VI	9	345	69%	57%
All Job Families	43	2,420	65%	46%

It should be emphasized again that the 43 jobs included in this study had not been used in any way in establishing the visual job family profiles. The profiles used (shown in Fig. 5.19) had already been established before the data on the 43 jobs were obtained. These jobs therefore represent strictly "hold-out" groups and the present investigation was a cross-validation of the previously established job family visual skill profiles.

The degree of discrimination among the 43 jobs varied a great deal, as would be expected, such variations being in part due to the relative importance of visual skills for the jobs in question. In general, however, the findings present very persuasive evidence that there is a definite relationship between vision test scores and success on many industrial jobs.

TESTS OF SPECIAL APTITUDES

Aside from the types of aptitude tests described above, there are many others that deal with other special aptitudes. We shall briefly mention a few of these.

Clerical Aptitude Tests. Several tests of clerical aptitudes have been developed. Although such tests usually measure certain aspects of mental abilities, they typically measure performance on certain specific tasks commonly performed in office work, which makes them especially useful in the selection of clerical workers.

One such test is the Minnesota Clerical Test,[92] which is a test of speed and accuracy in checking pairs of names and numbers. Another one is the Purdue Clerical Adaptability Test.[93] This test has several subtests as follows:

1. Spelling.
2. Computation.
3. Checking.
4. Word meaning.
5. Copying.
6. Arithmetical reasoning.

Perception Tests. In a psychological sense, *perception* refers to the ability to identify, recognize, discriminate, or otherwise attach meaning to various stimuli, such as things we see or sounds we hear. Perception, then, involves not only the sensory skill (such as vision or hearing), but *also* the mental recognition or discrimination of that which is seen or heard. Visual perception is especially important in some kinds of industrial jobs, such as in inspection operations. Although visual perception was identified by Thurstone[94] as a primary "mental" ability, we think it more appropriate to consider it in our present context of special aptitudes.

---

[92] Distributed by the Psychological Corporation.

[93] Distributed by the University Book Store, West Lafayette, Indiana.

A test of such ability is the Inspection test (test No. 1) of the FACT series,[95] previously mentioned. This test is intended to measure the ability to spot flaws or imperfections in objects.

VALIDITY OF APTITUDE TESTS FOR VARIOUS OCCUPATIONS

In the preceding sections of this chapter various studies have been mentioned that deal with the validity of aptitude tests for specific jobs. It may now be in order to present an over-view of the validity of different kinds of aptitude tests for the selection of people for various occupational groups. Two studies by Ghiselli and Brown[96] summarize the results of many published and unpublished test-validation studies in a manner that provides us with such an over-view. The first summary deals with studies relating to the prediction of performance in occupational training courses and of apprentices in training. The second summary deals with studies relating to the prediction of job proficiency. The second summary is particularly pertinent; it will be described briefly.

For each test-validation study that was included, information was available on (1) the job, (2) the test or tests used, (3) the criterion, and (4) the correlation between test scores and criterion measures (test validity). All of the specific tests used were originally classified into 16 different categories. Subsequently, on the basis of an analysis of the relationships among the tests, the tests were grouped into the following three "clusters":

---

[94] L. L. Thurstone, *Primary Mental Abilities, op. cit.*

[95] Distributed by Science Research Associates.

[96] E. E. Ghiselli and C. W. Brown, "Validity of Aptitude Tests for Predicting Trainability of Workers," *Personnel Psychology*, 1951, **4**, 243–260; C. W. Brown and E. E. Ghiselli, "Some Generalizations Concerning the Validity of Aptitude Tests," *Personnel Psychology*, 1953, **6**, 139–150.

A. *Intellectual ability* (including tests of intelligence, memory, substitution, arithmetic, number comparison, name comparison, and cancellation).

B. *Spatial ability* (including tests of tracing, location, pursuit, and spatial relations).

C. *Motor ability* (including tests of tapping, dotting, and finger, hand, and arm dexterity).

The specific jobs were then classified into 21 job groups, as shown in Fig. 5.20.

Then, by a method of analysis that need not be described here, the "relative validity" of each of the three types of tests was determined for each of 21 job categories. These are expressed in terms of an arbitrary scale,[97] which *cannot* be interpreted as a correlation. In other words, this figure simply shows an estimate of the *relative* validities of three types of tests for the various job classes.

It will be seen that intellectual ability had the highest relative validity for general supervisors, foremen, the clerical groups, salesmen, protective-service workers, and complex-machine operators. This ability has little relationship to success on such jobs as salesclerks, service workers, and vehicle operators.

The spatial ability tests showed the highest relative validity indexes for structural workers (which would be expected), and for some jobs for which such indexes would not be expected, namely computing and sales clerks, service workers, and gross manual workers. The validity indexes of spatial ability were lowest for general supervisors, salesmen, protective workers, and general and recording clerks. Most of the mechanical jobs have intermediate validity.

The motor ability tests also revealed some unexpected results. The high values are logical for such job groups as vehicle operators, processing workers, and possibly gross manual workers, but are not as easily explained for protective workers and salesclerks. For most of the industrial jobs, the validity indexes were of intermediate values.

Some of the unexpectedly low validity indexes (such as for the spatial ability test cluster for mechanical jobs) may be explained in part by "negative selection" within the groups covered by the study. The voluntary or involuntary elimination of some workers on a job (particularly the poorer workers) frequently tends to reduce the

---

[97] The scale is a standard score scale with a mean of 50 and a standard deviation of 20.

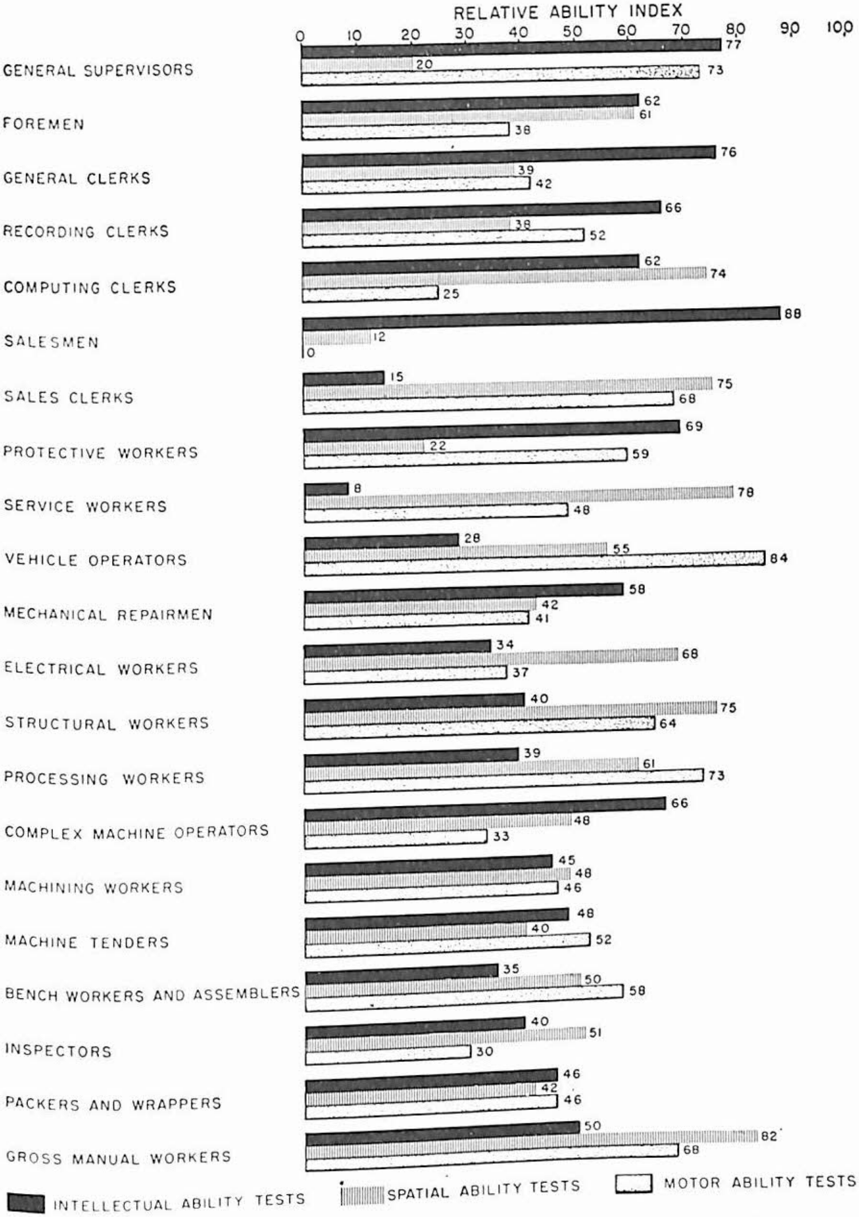


FIG. 5.20. "Relative validity" index of three types of tests for various classes of jobs. (Adapted from Brown and Ghiselli, op. cit.)

range of test scores, and also to reduce, spuriously, the correlation of the test with the criterion. This effect, in turn, may have caused the *relative* validities for other job groups to be somewhat higher than they really should be.

Although the results summarized need to be reviewed with certain reservations (such as the one mentioned above), it nonetheless helps to summarize a great deal of the work that has been carried out in validating aptitude tests. In particular, it can give cues regarding the relative importance of various types of aptitude tests for the selection of people for each of several job groups.

6

TESTS OF PERSONALITY AND INTEREST

Employment managers universally recognize the importance of personality traits in employees whom they hire. Indeed, one of their reasons for sometimes being hesitant in adopting psychological tests is that they often think of tests only in terms of intelligence or dexterity, and these tests do not, of course, take into account the more general personality traits of the applicant. An applicant might be very high in mental ability or in manipulative dexterity and yet have a personality that would not only unfit him for the job for which he is applying but would also make him a definitely undesirable individual to employ in any capacity.

The foregoing sections, which have dealt with aptitude and achievement tests without regard to the applicant's general personality, do not imply that the more general personality traits are unimportant. Psychologists are the first to recognize the importance of personality traits in helping an employee adapt himself to any job or to any organization. Psychological tests have emphasized specific aptitudes because psychologists recognize the importance of job aptitude, as such, aside from personality characteristics, and because, up to the present time, it has proved possible to develop adequate test of such aptitudes as finger dexterity or intelligence to a greater extent than it has been possible to develop adequate tests in the complicated field of personality.

Within recent years, however, both employment managers and consulting psychologists have increasingly demanded some reasonably satisfactory and accurate method of determining certain personality traits of an applicant at the time he applies for employment. It is now recognized not only that an applicant who does not have the aptitude to learn the job will fail, no matter how desirable his personality traits may be, but also that if he does have the aptitude for the job he will probably still fail if his personality characteristics make it difficult for him to fit into the organization and to work cooperatively with other persons. Although this conclusion is sound common sense, we do not need to rely entirely upon subjective judgment to reach it. Such investigations as those of Hunt[1] and Brewer,[2] which have been carried on with thousands of employees in a variety of industries, show that personality factors, rather than lack of ability on the job, are responsible for a large number of layoffs and failures to be promoted. These studies deal with layoffs due to factors other than failure to produce. The rising voice of labor in determining managerial policies, particularly with respect to hiring and firing, and the growing importance of seniority as a determining factor in layoff, have made it increasingly important to determine, at the time of hiring, whether an applicant has any incipient personality maladjustments that might prevent him from fitting properly into the organization.

There is also general recognition among employment managers of the importance of the interests of employees in their work. Comparing persons on similar jobs, there is reason to believe that those who do *not* like their jobs are not as good in job performance, and do not stay as long, as those who *do* like their jobs. Moreover, the individuals who are doing work in which they have little or no interest do not have the opportunity to achieve the level of job satisfaction that is experienced by persons who are doing what they like to do.

It has been with the hope of measuring personality traits and interests that psychologists have developed various personality and interest tests. In the development and use of such tests, however, difficulties have been encountered which presently place serious limitations on their use for employment purposes. In particular, these limitations stem from the possibility of "faking" on the test, that is, the possibility of an applicant giving responses to test items that are not descriptive of himself. More will be said about this later, but first it would be desirable to describe briefly some of the personality and interest tests that are available.

---

[1] H. C. Hunt, "Why People Lose Their Jobs or Aren't Promoted," *Personnel Journal,* 1935, **14,** 227.

[2] J. M. Brewer, "Religion and Vocational Success," *Religious Education,* 1930, **25,** 29–41.

TESTS OF PERSONALITY

Most of the personality tests being used in industry are paper-and-pencil tests, with multiple-choice items. The items typically consist of statements, or questions, relating to behavior, attitudes, or beliefs. A typical item is given below:

1. I feel uncomfortable with other people.
    a) Yes
    b) Don't know
    c) No

The responses typically are scored to give a score on each of several personality or temperament "traits," such as "dominance," or "sociability."

In a few companies "projective" tests have been used. These usually are individual (rather than group) tests. In a projective-type test the subject is presented with each of several "items" such as ink blots or pictures, and he is asked to tell what he "sees" in the object in question. A trained test administrator evaluates his responses and makes an assessment of his personality.

The Bernreuter Personality Inventory.[3] This is one of the most widely used personality tests. It consists of 125 questions dealing with a wide variety of topics that indicate personality traits and characteristics. The scale may be scored with each of six different scoring keys, each of which measures a different personality component. The components measured, with the codes used in referring to them, are:

B1-N. A measure of neurotic tendency. Persons scoring high on this scale tend to be emotionally unstable. Those scoring above the 98 percentile would probably benefit from psychiatric or medical advice. Those scoring low tend to be very well balanced emotionally.

B2-S. A measure of self-sufficiency. Persons scoring high on this scale prefer to be alone, rarely ask for sympathy or encouragement, and tend to ignore the advice of others. Those scoring low dislike solitude and often seek advice and encouragement.

---

[3] R. G. Bernreuter, *The Personality Inventory* and *Manual* (Stanford, Calif., Stanford University Press, 1935).

B3-I. A measure of introversion-extroversion. Persons scoring high on this scale tend to be introverted; that is, they are imaginative and tend to live within themselves. Scores above the 98 percentile bear the same significance as do similar scores on the *B1-N* scale. Those scoring low are extroverted; that is, they rarely worry, seldom suffer emotional upsets, and rarely substitute day-dreaming for action.

B4-D. A measure of dominance-submission. Persons scoring high on this scale tend to dominate others in face-to-face situations. Those scoring low tend to be submissive.

F1-C. A measure of confidence in oneself. Persons scoring high on this scale tend to be hamperingly self-conscious and to have feelings of inferiority; those scoring above the 98 percentile would probably benefit from psychiatric or medical advice. Those scoring low tend to be wholesomely self-confident and to be very well adjusted to their environment.

F2-S. A measure of sociability. Persons scoring high on this scale tend to be nonsocial, solitary, or independent. Those scoring low tend to be sociable and gregarious.

Scoring keys for the first four of the components (indicated by code *B*) were developed by Bernreuter[4] in his original work with the scale. The last two (indicated by code *F*) were derived from a factor analysis of the Bernreuter scale reported by Flanagan.[5]

Because of the long time required to score the Bernreuter test with the original scoring keys, a simplified scoring system has been developed and reported by McClelland.[6] The simplified scoring system proposed gives scores that, for components *B1-N*, *B4-D*, *F1-C*, and *F2-S*, correlate above .95 with scores obtained by the original scoring method. For component *B2-S* the correlation is .84. A simplified scoring for component *B3-I* was not developed because the latter correlates so highly with *B1-N*.[7]

---

[4] R. G. Bernreuter, "The Theory and Construction of the Personality Inventory," *Journal of Social Psychology*, 1933, **4**, 387–405.

[5] J. C. Flanagan, *Factor Analysis in the Study of Personality* (Stanford, Calif., Stanford University Press, 1935).

[6] D. C. McClelland, "Simplified Scoring of the Bernreuter Personality Inventory," *Journal of Applied Psychology*, 1944, **28**, 414–419.

Several investigations conducted in business and industrial situations have been reported that show the effectiveness of the Bernreuter Personality Inventory. Richardson and Hanawalt[8] found that the Bernreuter scales differentiate between "office holders" (men who have had at least two presidencies or important chairmanships since the age of twenty-one in business, professional, civic, religious, fraternal, or social organizations) and men who have not held such offices.

The "office holders" made scores that showed them to be significantly more dominant, more self-confident, less neurotic, and less introverted than those who had held no office. In this same study, Richardson and Hanawalt reported significant differences between the Bernreuter scores of supervisors (men who had fifteen or more persons under their direction or supervision) and those of a contrasting nonsupervisory group. The supervisors made scores showing them to be less neurotic, less introverted, more dominant, and more self-confident than the non-supervisory group.

The Humm-Wadsworth Temperament Scale.[9] This scale has been developed by two men whose work has been primarily in the field of industrial personnel. The scale consists of 318 questions, which the person taking the test answers by checking *yes* or *no*. From the answers to these questions it is possible, by differential scoring, to obtain separate scores for seven aspects of temperament. These aspects are as follows:

> *The Normal Component.* This is primarily a control mechanism providing rational balance and temperamental equilibrium. It underlies the conservatism, toleration, and conformity to socially acceptable behavior observed in the well-adjusted person.
>
> *The Hysteroid Component.* An individual with an excess of the hysteroid component possesses a character defect with ethically inferior motiva-

---

[7] R. G. Bernreuter, *Manual for the Personality Inventory* (Stanford, Calif., Stanford University Press, 1935).

[8] H. M. Richardson and N. G. Hanawalt, "Leadership as Related to the Bernreuter Personality Measures: III. Leadership Among Adult Men in Vocational and Social Activities," *Journal of Applied Psychology*, 1944, **28**, 308–317.

[9] D. G. Humm and G. W. Wadsworth, *The Humm-Wadsworth Temperament Scale, Test Booklet and Manual* (2nd rev.; Doncaster G. Humm Personnel Service, Los Angeles, California, 1940).

tion manifested by stealing, lying, cheating, and similar antisocial behavior.

The Manic Cycloid Component. This is characterized by emotionality, fluctuation in activities, and interferences with voluntary attention, some degree of elation, pressure of activity, and distractibility together with such manifestations of excitement as jests, pranks, enthusiasms, impatience, and so forth.

The Depressive Cycloid Component. This is manifested by some degree of sadness, lessened activity, and associated characteristics such as worry, timidity, and feeling of malaise.

The Autistic Schizoid Component. This is characterized by heightened imagination, leading to a tendency toward a daydream life concerning which the subject is sensitive.

The Paranoid Schizoid Component. This includes stubborn adherence to fixed ideas, suspicion, and contempt for the opinion of others, with behavior fitting these traits.

The Epileptoid Component. This is characterized by inspirations to achievement that are meticulously developed and pushed through to completion.

Everyone who studies the above list of components with their brief descriptions will recognize certain persons of his acquaintance who, from long observation, are clearly known to possess an excessive amount of one or another or of some combination of these characteristics. One does not need the Humm-Wadsworth Scale, or perhaps even training in psychology, to identify the Hysteroid or the Cycloid if he is constantly thrown into contact with an individual of this type. Indeed, one thrown into contact with such a person is more likely to recognize that something is wrong than is the afflicted individual himself. The mentally ill or "near ill" often feel that they are quite normal but that everybody else is wrong. It is, however, one thing to recognize the presence of such personality characteristics in a person with whom one is thrown into constant contact, and it is quite another thing to recognize them in an applicant whom one is interviewing for the first time. It is claimed that the Humm-Wadsworth Temperament Scale will accomplish this result.

The Guilford Series of Personality Tests. Guilford and his collaborators[10] have conducted an extensive series of statistical investigations that have identified a number of personality factors. These investigators have also devised scales to measure several of these factors identified. The Guilford series of three personality scales follows:

---

[10] J. P. and R. B. Guilford, "Personality Factors S, E, and M, and their Measurement," *Journal of Psychology*, 1936, 2, 107–127; "Personality Factors D, R, T, and A," *Journal of Abnormal and Social Psychology*, 1939, 34, 21–36; and "Personality Factors N and GD," *Journal of Abnormal and Social Psychology*, 1939, 34, 239–248.

An Inventory of Factors S T D C R[11]

This scale measures the factors:

S—Social introversion-extroversion. Shyness, seclusiveness, tendency to withdraw from social contacts, versus sociability, tendency to seek social contacts, and to enjoy the company of others.

T—Thinking introversion-extroversion. An inclination to meditative or reflective thinking, philosophizing, analysis of one's self and others, versus an extrovertive orientation of thinking.

D—Depression. Habitually gloomy, pessimistic mood, with feelings of guilt and unworthiness, versus cheerfulness and optimism.

C—Cycloid disposition. Strong emotional fluctuations, tendencies toward flightiness and emotional instability, versus uniformity and stability of moods, evenness of disposition.

R—Rhathymia. A happy-go-lucky, carefree disposition, liveliness, impulsiveness, versus an inhibited, over-controlled, conscientious, serious-minded disposition.

Guilford-Martin Inventory of Factors G A M I N[12]

This scale measures the factors:

G—General pressure for overt activity.

A—Ascendancy in social situations as opposed to submissiveness; leadership qualities.

M—Masculinity of attitudes and interests as opposed to femininity.

I—Lack of inferiority feelings; self-confidence.

N—Lack of nervous tenseness and irritability.

The Guilford-Martin Personnel Inventory[13]

This scale measures the factors:

O—Objectivity (as opposed to personal reference or a tendency to take things personally).

Ag—Agreeableness (as opposed to belligerence or a dominating disposition and an overreadiness to fight over trifles).

Co—Cooperativeness (as opposed to faultfinding or overcriticalness of people and things).

The three components of this inventory all measure certain aspects of the paranoid trait of temperament. Since an excessive amount of this trait has long been considered to underlie many of the difficulties encountered by some people in situations requiring contact with others, this scale would seem to have ideal possibilities for use in industrial situations. An indication of its validity in an industrial situation is given by the results obtained in a Southern California company.[14] The test was given to 51 employees, among whom were 22 who were regarded by management as trouble-makers or malcontents. The scores for traits *O*, *Ag*, and *Co* yielded very flat distributions for the 51 cases. Having no knowledge of the proportion of satisfactory and unsatisfactory workers in the group, and having no knowledge of which individuals belonged in each category, it was decided, in view of the form of the test score distributions, to accept the medians as the critical scores. Any individual who was below the median on two or more traits was predicted to be in the unsatisfactory group. Of the 22 workers who had been judged unsatisfactory by management, 73 per cent were detected as unsatisfactory in terms of the critical scores used. Of the 29 who had been judged satisfactory by management, only 34 per cent were mistakenly placed in the unsatisfactory category. Guilford suggests that a test that makes more than twice as many correct predictions as errors can be said to offer definite usefulness.

---

[11] Distributed by Sheridan Supply Co., Beverly Hills, California.

[12] *Ibid.*

[13] *Ibid.*

The Minnesota Multiphasic Personality Inventory.[15] This inventory consists of 550 statements that the person being tested sorts into three piles—those he regards as *true*, those he regards as *false*, and those on which he cannot make a true or false judgment. After the cards have been sorted, the scoring procedure yields eight different personality component scores, viz:

---

[14] *The Guilford-Martin Personnel Inventory, Manual of Directions and Norms* (Sheridan Supply Co., Beverly Hills, California).

[15] Distributed by the Psychological Corporation. The research on which the scales are based is reported in the following articles:

S. R. Hathaway and J. C. McKinley, "A Multiphasic Personality Schedule (Minnesota): I. Construction of the Schedule," *Journal of Psychology*, 1940, **10**, 249–254; "**II.** A Differential Study of Hypochondriasis," *Journal of Psychology*, 1940, **10**, 255–268; "**III.** The Measurement of Symptomatic Depression," *Journal of Psychology*, 1942, **14**, 73–84; "IV. Psychasthenia," *Journal of Applied Psychology*, 1942, **26**, 614–624.

J. C. McKinley and S. R. Hathaway, "The Minnesota Multiphasic Personality Inventory: V. Hysteria, Hypomania, and Psychopathic Deviate," *Journal of Applied Psychology*, 1944, **28**, 153–174.

Hypochondriasis Scale. Measures undue worry about one's health.

Depression Scale. Measures feeling of unworthiness, uselessness, and undue worry.

Hysteria Scale. Measures the degree to which subject is like patients who have general systemic complaints and/or specific complaints such as paralyses, contractures, and gastric, intestinal, and/or cardiac symptoms.

Psychopathic Deviate Scale. Measures similarity of subject to patients whose main difficulty lies in their absence of deep emotional response, inability to profit from experience, and disregard of social mores.

Interest Scale. Measures tendency toward masculinity or femininity of one's interest pattern.

Paranoia Scale. Measures stubborn adherence to fixed ideas, suspicion, and contempt for the opinion of others.

Psychasthenia Scale. Measures similarity of the subject to psychiatric patients who are troubled with phobias (unreasonable fears) or compulsions.

Schizophrenia Scale. Measures similarity of subject to psychiatric patients who are characterized by bizarre and unusual thoughts and behavior.

In addition to the eight personality components measured, the Minnesota Multiphasic Personality Inventory also provides for certain checks on the accuracy and honesty of the subject's responses. A further expansion of the scoring possibilities of the scale has been developed by Drake[16] in the form of a key to obtain scores on social introversion-extroversion. Simplified scoring of the scale for group administration, to be used when electric, punched-card tabulating equipment is available, has been described by Burton and Bright.[17]

An indication of the validity of the Minnesota Inventory comes from the work of Schmidt,[18] who reports that the Inventory "distinguished . . . much statistical significance between normal soldiers and those diagnosed as constitutional psychopaths; mild or severe neurosis; and psychosis."

The Thurstone Temperament Schedule.[19] This schedule is specifically designed to survey those relatively stable personality traits important for success and satisfaction in different types of work. Seven different traits are covered, namely: *active, vigorous, impulsive, dominant, stable, sociable, and reflective.*

---

[16] L. E. Drake, "A Social I. E. Scale for the Minnesota Personality Inventory," *Journal of Applied Psychology*, 1946, **30**, 51–54.

[17] A. Burton and C. J. Bright, "Minnesota Multiphasic Personality Inventory for Group Administration," *Journal of Consulting Psychology*, 1946, **10**, 99–103.

[18] H. O. Schmidt, "Test Profiles as a Diagnostic Aid: The Minnesota Multiphasic Inventory," *Journal of Applied Psychology*, 1945, **29**, 115–131.

[19] L. L. Thurstone, *Thurstone Temperament Schedule* (Chicago, Science Research Associates, 1951).

Gordon Personal Profile and Personal Inventory.[20] These two relatively new tests have what are called "forced-choice" items. Each item includes four statements, such as the following:

Prefers to get up early in the morning.
Doesn't care for popular music.
Has an excellent command of English.
Obtains a poorly balanced diet.

The person taking the test is asked to indicate which of these statements is *most* like himself, and which is *least* like himself. The forced-choice technique will be discussed more in detail in Chapter 8. It might be said here, however, that the four statements in each group have been subjected to previous analyses, and have been formed into these groups on the basis of those analyses. In general, they have been so selected that it is difficult for the person being tested to figure out what would be the "best" answer to give in case he wants to "fake" the test. Thus, tests built in this manner are considered to be less susceptible to the possibility of faking.

The tests are scored on various scales, as indicated below:

Gordon Personal Profile
A – Ascendancy
R – Responsibility
E – Emotional Stability
S – Sociability
T – Total Score

Gordon Personal Inventory
C – Cautiousness
O – Original Thinking
P – Personal Relations
V – Vigor
T – Total Score

The Rorschach Projective Test. There have been several attempts, either unsuccessful or incomplete, to use the Rorschach projective test in the selection of industrial personnel. The Rorschach is a projective test—one in which the person tested "projects" his own interpretation into certain standardized stimulus situations that are presented to him. The stimuli in the case of the Rorschach test are standardized ink blots. Kurtz[21] reports an attempt to validate this test in the selection of life insurance sales managers. The responses of successful and unsuccessful groups of managers were item-analyzed to derive a special scoring key to differentiate between the two groups. The key thus derived correctly classified 79 of the 80 individuals comprising the two original groups. The necessity of cross-validating a key of this sort[22] on other groups, which were not used in establishing the key, was shown by the fact that when the key was tried with different groups, it did not differentiate better than could be accounted for by chance. A similar study of the Rorschach by Cox[23] gave results on original groups that suggest the usefulness of this test for the selection of salesclerks. The necessary step of cross-validation on new criterion groups, however, was not reported in this study.

---

[20] Distributed by the World Book Co., Yonkers-on-Hudson, New York.

INTEREST TESTS

Interest tests typically require the persons being tested to indicate the strength of his interests in such things as various jobs, hobbies, recreation, or leisure-time activities. This is sometimes done by presenting groups of, say, three activities, and by asking the individual which he likes most, and which he likes least. In other cases, he simply indicates, for each statement, the extent to which he likes or dislikes the activity in question.

Strong Vocational Interest Blank.[24] This test has been developed primarily for use in vocational guidance counseling. It determines whether the subject's pattern of interests agrees with the interest pattern of men in each of a large number of professions and occupations. Other things being equal, a person choosing a profession or occupation is more likely to be happy and successful if his basic interests are similar to the interests of men actually in this particular field.

---

[21] A. K. Kurtz, "A Research Test of the Rorschach Test," *Personnel Psychology*, 1948, **1**, 41–51.

[22] E. E. Cureton, "Validity, Reliability, and Baloney," *Educational and Psychological Measurement*, 1950, **10**, 94–96.

[23] K. J. Cox, "Can the Rorschach Pick Sales Clerks?" *Personnel Psychology*, 1948, **1**, 357–363.

[24] E. K. Strong, Jr., "Vocational Interest Test," *Educational Record*, 1927, **8**, 107–121. This test is distributed by the Psychological Corporation.

Different forms of this test and different scoring keys are available, so that it may be used with men or women of student or adult age.

The Strong Vocational Interest Blank measures as many patterns of interest as vocational patterns for which it is scored. This procedure is advisable in vocational guidance work, and the test is admirably suited for that purpose. Recent work, however, has indicated that interests may be grouped into patterns, the number of which is considerably smaller than the possible number of vocations from which a choice must be made. Thurstone[25] reports a multiple factor study of vocational interests that shows by an analysis of Strong-Blank results that interests may be basically divided into four general fields: science, language, people, and business. It would seem to be important for the personnel manager to consider these interests as well as abilities, particularly in the case of college graduates or others who are expected to show considerable development over a period of years. It is probably unwise to place in the scientific development department young college men whose basic interests are in people and business; and it would seem to be equally unwise to put into the sales department those whose basic interests are scientific. Fortunately, it is often possible to obtain a fairly accurate picture of an applicant's or employee's interests by means of an interview in which a number of questions are asked about such topics as present and past activities, hobbies, and how vacations are spent.

Kuder Preference Record.[26]

There are four forms of the Kuder Preference Record, two of which will be described. Form C (Vocational) measures ten broad areas of interests, as follows:

Outdoor
Mechanical
Computational
Scientific
Persuasive

---

[25] L. L. Thurstone, "A Multiple Factor Study of Vocational Interests," *Personnel Journal*, 1931, **10**, 198–205.

[26] G. F. Kuder, "The Stability of Preference Items," *Journal of Social Psychology*, 1939, **19**, 47–50. This test is distributed by Science Research Associates, Chicago, Illinois.

Artistic
Literary
Musical
Social Service
Clerical

Separate scores are readily obtained for each of the listed interest groups. This type of scoring makes the test particularly easy to try out on a "test the test" basis (see pp. 81 ff.) because only the ten basic scores will be considered, and these yield an over-all pattern of interest. The manual with this test, however, provides interpretive data for evaluating various interests patterns in terms of their agreement with those of men in various occupations.

Form D (Occupational) measures the individual's interest in certain specific occupations, rather than in broad vocational areas. This is a new form of the test, published in 1956. At present, separate answer keys are available for scoring a person's interests in each of ten occupations, as follows:

Electrical Engineer
Farmer
Forester
Minister
Newspaper Editor
Physican
Psychologist—Clinical
Psychologist—Industrial
School Superintendent
YMCA Secretary

Just as in the case of personality tests of the type previously discussed, the scores on an interest test, such as the Kuder Preference Record, may be distorted by an applicant or employee who attempts to "fake" his answers. Kuder[27] has described two methods of scoring the Preference Record to differentiate between blanks answered sincerely and blanks answered with the purpose of making a good impression. These scores were found to identify faked *vs.* sincerely answered blanks with from 87 to 94 per cent accuracy.

---

[27] G. F. Kuder, "Identifying the Faker," *Personnel Psychology*, 1950, 3, 155–167.

Cardall Primary Business Interests Test.[28] This test is designed to measure an individual's preferences for the specific job activities that characterize beginning business jobs. These immediate and specific preferences point to the initial job, determine the individual's interest or boredom in his first activities, and often determine to a considerable extent his progress in his work. The five business fields in which this blank measures interest are: accounting, collecting and adjusting, sales office work, sales store work, and stenographer-filing.

Several investigations have indicated that the interest questionnaires, when used in this manner, are of value in industry. Shartle[29] made a study of two groups of foremen, one consisting of highly successful supervisors and the other made up of foremen approximately equal in job skill but less able than the first group in handling the personal aspects of their supervisory jobs. In Shartle's investigation the Strong Vocational Interest Blank was given to all supervisors, and this was followed by a detailed clinical interview. The results showed that the less-successful foremen were characterized by more interest in withdrawing from others, by more indifference to the actions of others, and by more antagonistic reactions toward others. The study suggests that interest in dealing with others is one of the prime requisites of the successful supervisor.

LIMITATIONS OF PERSONALITY AND INTEREST TESTS

As implied above, personality and interest tests have serious limitations for employment purposes, the most important one being the possibility of faking on the part of applicants.

Faking of Personality and Interest Tests. If a person is taking a personality or interest test for the purpose of receiving vocational guidance, he will usually be motivated to give relatively truthful answers to the test questions, since it is to his interest to find out all he can about himself in order to make a sound vocational choice. If he is applying for a job, however, his motivation to get the job may induce him to give responses that he *thinks* will make him out to be the kind of person the company is looking for.

---

[28] A. J. Cardall, "A Test for Primary Business Interests Based on a Functional Occupational Classification," *Educational and Psychological Measurements*, 1942, 2, 113–138. This test is distributed by Science Research Associates, Chicago, Illinois.

[29] C. L. Shartle, "A Personnel Approach to Foremanship," *Personnel Journal*, 1943, 13, 135–139.

Evidence of the ability of people to fake such tests is available from various studies. One such study, for example, was carried out with the Humm-Wadsworth Temperament Scale with 65 college students.[30] Each student was given the scale twice: first, in a "clinical" situation, with instructions to be as frank as possible, and second, in an "employment" situation in which each student was to assume that he was in an employment office after a job and had been asked to take the test as a part of the employment procedure. Table 6.1 shows the mean scores for the seven components obtained under these two conditions.

TABLE 6.1

Mean Scores on Humm-Wadsworth Temperament Scale Obtained in a Clinical and an Employment Situation by 65 College Students*

Component	*Clinical Mean*	*Employment Mean*	*Shift from Clinical to Employment*
Normal	981	1023	+42
Hysteroid	1023	980	−63
Manic	1035	937	−98
Depressive	1061	913	−148
Autistic	1024	938	−86
Paranoid	970	955	−15
Epileptoid	983	1002	+19

* All scores were computed by the log method, with correction for no-count, as described in the second 1940 revision of the manual of directions.

It is apparent from Table 6.1 that the employment situation, when compared with the clinical, shows a higher average value for the normal component and lower values for all except the epileptoid of the remaining six components. In other words, the students were able, by assuming an attitude of "applying for a job," to change their test profiles toward more of the normal and less of the undesirable traits. All differences shown in Table 6.1 are significant from a statistical viewpoint, and only those scores were used in the computation that fell within the no-count limits, for both the clinical and employment situations, within which the manual of instructions states that the scale should be "accepted as probably valid."

Even with the shift in means from one situation to the other, it would still be possible to infer one's score in one situation from a knowledge of his score in the other situation if scores in the two situations were highly correlated. The correlations were computed and are shown in Table 6.2.

---

[30] This study was conducted by W. J. Giese and F. C. Christy at Purdue University.

TABLE 6.2

Correlations between Clinical and Employment Situations on the Humm-Wadsworth Temperament Scale for 65 College Students

Component	*Correlation*
Normal	—.03
Hysteroid	+.42
Manic	+.09
Depressive	—.10
Autistic	+.11
Paranoid	+.61
Epileptoid	+.23

The only correlation in Table 6.2 that is large enough to be significant is the one for the paranoid component.

Without going into any details, it might be said that other somewhat corresponding studies with other tests have given similar results. For example, Longstaff[31] found that both the Strong Vocational Interest Blank and the Kuder Preference Record could be faked to a considerable degree, although some interest categories were more fakable than others. In another study, Longstaff and Jurgensen[32] found that scores on the Jurgensen Classification Inventory could be changed significantly, depending on the mental "set" of the individuals at the time as affected by the instructions given to the students taking the test.

The fact that people can fake personality and interest inventories implies serious limitations to the potential users of such tests for employment purposes. There is, however, the possibility that tests can be developed that are not as fakable as some of the current tests. The most promising development along these lines is the use of forced-choice tests.

Fakability of Forced-Choice Tests. The Gordon Personal Profile and the Gordon Personal Inventory, previously mentioned, are based on the so-called forced-choice technique. By this method, the "favorable" items within a group have been determined to be equally favorable, and the "unfavorable" items within a group have been determined to be equally unfavorable. Because of this, a person cannot choose, as the statement *most* like him, the *most* favorable answer, because there will be *two* answers that appear to be *equally* favorable. Similarly, in selecting a statement that is *least* like him, he has to choose from two that seem to be *equally* unfavorable. This structure is intended to reduce the possibility of faking, and there is some evidence that it does.

---

[31] H. P. Longstaff, "Fakability of the Strong Interest Blank and the Kuder Preference Record," *Journal of Applied Psychology*, 1948, 32, 360–369.

[32] H. P. Longstaff and C. E. Jurgensen, "Fakability of the Jurgensen Classification Inventory," *Journal of Applied Psychology*, 1953, 37, 86–99.

This evidence comes from two studies with the Gordon Personal Profile, one by Rusmore[33] and the other by Gordon and Stapleton.[34] In the first of these studies the same group of 81 college students took the test under a simulated "industrial" situation and under a simulated "guidance" situation. In the second study 121 high school students first took the test in connection with the regular guidance program, and three months later took the test when actually applying for outside summer employment. The results of these two studies are summarized in Table 6.3. This shows the means of the scores on the various test components for the two administrations for each group. In the case of the college students it can be seen that the simulated "industrial" situation resulted in a significant difference on the *Responsibility* and *Total* scores, although these differences were only about 9 and 8 per cent respectively. In the case of the high school students significant differences occurred in these two scores as well as in the *Emotional Stability* scale. The difference in the case of the *Responsibility* scale is rather substantial; the differences in the *Emotional Stability* and *Total* scores are more moderate (although statistically significant). These moderate differences, in combination with the fact that the differences for the remaining two scales were not significant suggest that forced-choice personality tests may be somewhat more foolproof against faking than the more conventional personality tests.

It must be recognized that the results of these studies of college and high school students cannot be generalized to industrial situations. It

---

[33] J. T. Rusmore, "Fakability of the Gordon Personal Profile," *Journal of Applied Psychology*, 1956, **40**, 175–177.

[34] L. V. Gordon and E. S. Stapleton, "Fakability of a Forced-Choice Personality Test Under Realistic High School Employment Conditions," *Journal of Applied Psychology*, 1956, **40**, 258–262.

TABLE 6.3

Mean Scores on the Gordon Personal Profile Taken Twice Under Specified Test Conditions

	GROUP AND CONDITION			
	College Students (81)		H. S. Students (121)	
TEST COMPONENT	*Simulated "Guidance"*	*Simulated "Industrial"*	*Guidance*	*Employment*
Ascendancy	3.7	4.4	2.9	3.2
Responsibility	7.4	8.9 †	4.3	6.8 †
Emotional Stability	7.3	8.0	5.3	6.8 †
Sociability	4.4	4.8	4.8	4.5
Total	23.0	26.1*	17.2	21.3 †

* Difference between the "Guidance" and "Industial" or "Employment" mean scores significant at the 5% level of confidence.

† Difference between the two means significant at the 1% level of confidence.

Source: Adapted from articles by Rusmore and by Gordon and Stapleton, *op. cit.*

remains to be determined whether forced-choice personality tests are less fakable in real industrial situation than are the more conventional tests. The results of these studies, however, suggest that forced-choice personality and interest tests may offer a partial solution to the present dilemma of personality and interest tests for employment purposes.

THE USE OF PERSONALITY AND INTEREST TESTS

There is a great temptation on the part of management to use personality and interest tests without taking necessary precautions in their interpretation. The following type of "reasoning" represents a fairly common mistake in adopting such tests for employment purposes:

1. One personality component measured by a given test is called *neurotic tendency*. This test *must* measure *neurotic tendency* because that title is used by the authors of the test. (The reader may substitute any presumably undesirable personality trait he wishes for the phrase *neurotic tendency*. The result of the "reasoning" will be the same.)
2. We do not want in our plant people who have a high degree of *neurotic tendency*.
3. Therefore we do not want in our plant anyone who scores beyond a certain point on this test.

The same "reasoning" can be applied to any other "undesirable" trait or interest, or can be applied in reverse to any "desirable" trait or interest.

Anyone familiar with the statistical procedures that should be followed in developing and standardizing any test for use in a given situation will recognize a fallacy in every one of the three steps given above. First, a test does not necessarily measure a given trait merely because the name of that trait is used in the title or description of the test. Second, one cannot assume that persons with *any* given trait, *as revealed by a score on a test,* are by virtue of that fact alone necessarily undesirable employees. It follows that the third statement, based upon the accepted truth of the first and second, is hardly justified without more specific evidence on the applicability of the test to the situation in question.

The Validation of Personality and Interest Tests. The validity of personality and interest tests for employment purposes must be determined by statistical analysis, in the same way as the validity of other tests must be determined. Because of the possible faking of such tests, however, one cannot *really* rely on their validity unless it has been determined by the follow-up method, and unless it has been double-checked by cross-validation on a second "hold-out" group.

Since the responses of applicants to such tests may be different from what their responses would be if they were later being tested as "present employees," it is important that such tests be validated with the *kinds* of responses that *applicants* typically give. It is possible that the validation of tests under such circumstances may result in little or no relationship with the criterion of job success. If this turns out to be the case, then it is better to face this reality than to use the test on the assumption that it is a useful device when it really is not. But it is also possible that the responses given by applicants *may* be related to the criterion. If this is the case, the test can be used for prediction whether the responses are, or are not, in all respects truthful responses.

A study by Tiffin and Phelan[35] illustrates this approach. A metal-parts factory had been giving the Kuder Preference Record (Vocational Form C-H) to applicants for hourly-paid jobs for a period of 14 months. The test results, however, had not been used in the employment process, and it was possible to conduct a follow-up study covering nearly 1,800 men who had been employed during that period. The analysis was made against a criterion of job tenure of the men who had been tested, 1,109 of whom were still employed at the time of the study. Of the remaining 684 who had left, 450 were "voluntary" quits. The tests of the 1,109 present employees and the 450 voluntary quits were randomly assigned to two groups. Two-thirds were used to form a "primary" group for item and subscore analysis; the remaining one-third were set aside as a "hold-out" group for later cross-validation. From the primary group two "extreme" criterion groups were selected; 250 present employees with 11–14 months' tenure formed a "long-tenure" group, and 200 voluntary quits, who had worked for three months or less, formed a "short-tenure" group.

---

[35] J. Tiffin and R. F. Phelan, "Use of the Kuder Preference Record to Predict Turnover in an Industrial Plant," *Personnel Psychology*, 1953, **6**, 195–204.

Two analyses were made for these two criterion groups. The first consisted of an item analysis; 76 items having a difference between the two groups of 10 per cent or more were retained to form a new scoring key. The second analysis was made on the basis of the correlations of scores on the several subtest components with the tenure criterion groups. Four of the scales gave evidence of a significant relationship with tenure. The "Outdoor" and "Mechanical" scales were positively correlated with tenure, while the "Persuasive" and "Clerical" scales were negatively correlated. These scales were incorporated in a composite score formula with appropriate, statistically determined weights.

Both the special scoring key and the composite scores (based on subtests) were then tried out with the hold-out group to see to what extent they would differentiate that group on the basis of tenure. For this purpose that group was divided into a "long-tenure" group (those who stayed more than three months) and a "short-tenure" group (those who stayed three months or less). The relationships are shown in Fig. 6.1. This shows, for the two methods of scoring (special scoring key and composite scores) the relationship between scores and tenure, specifically the per cent of short-tenure employees for each test-score group. Both scoring systems were found to discriminate to an appreciable degree. However, the special scoring key differentiated better, but it was not possible in this study to determine whether the difference between the two methods was statistically significant.

With this type of cross-validated evidence, a company could use the test in question with considerable confidence that it would really be useful in screening out before employment many persons who would most likely be short-tenure employees.

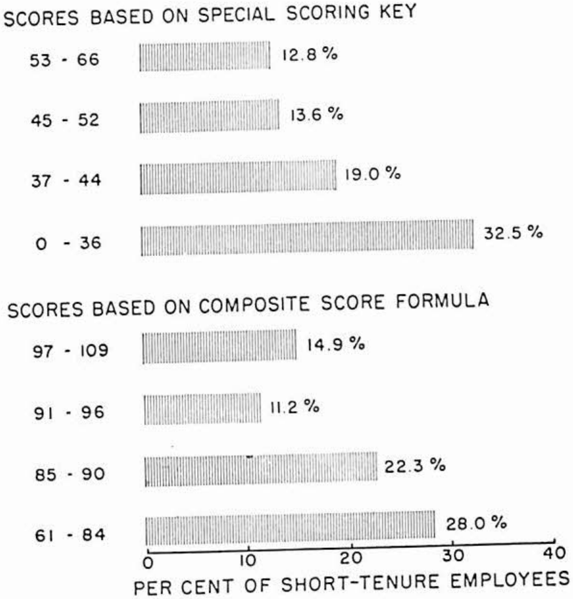


FIG. 6.1. Per cent of short tenure employees by test-score categories based on special scoring key and on composite score formula as used with Kuder Preference Record. (Adapted from Tiffin and Phelan, op. cit.)

Validity of Personality Tests for Employment. Before ending the discussion of personality tests, it would be desirable to indicate in some way the extent to which they have been shown to have validity for employment purposes. A survey by Ghiselli and Barthol[36] brings together, and summarizes very effectively, a total of 113 studies dealing with the validity of personality tests for employment purposes. About 40 per cent of the material in the survey was from unpublished sources. The studies included in the survey were restricted to those in which the criterion used was some index of job proficiency, such as production records or ratings by supervisors. An attempt was also made to include only those studies in which the scoring key for the personality test was developed independently of the group for which the validity coefficient was reported; no mention was made, however, of the extent to which the cases actually were limited to this type.

---

[36] E. E. Ghiselli and R. P. Barthol, "The Validity of Personality Inventories in the Selection of Employees," *Journal of Applied Psychology*, 1953, 37, 18–20.

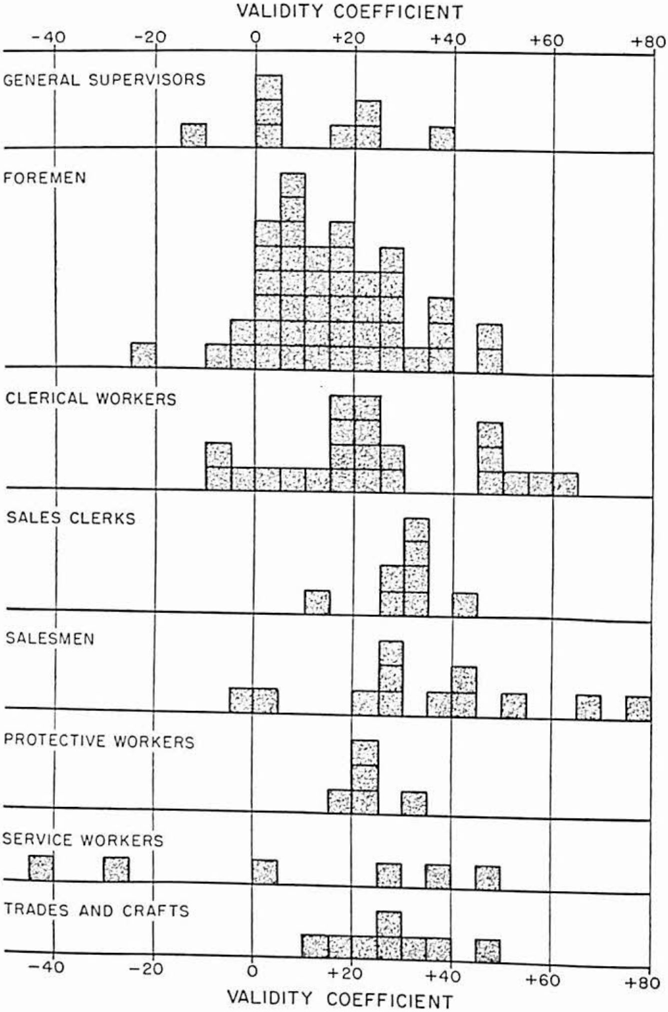


FIG. 6.2. Distribution of validity coefficients of personality inventories for various occupational groups. Decimals are omitted from coefficients (From Ghiselli and Barthol, op. cit.)

The results are shown in two ways. Fig. 6.2 gives the distribution of validity coefficients of the tests for various occupational groups, and Table 6.4 lists the weighted mean validity coefficients for the occupational groups. On the basis of these results, personality tests seem to be potentially more useful for selection of clerks, salesclerks, salesmen, protective workers, and trades and crafts, than for the other occupational groups. It can be seen, however, that in some groups the data are based on only a few cases, as with the protective workers and service workers. It can also be seen (from Fig. 6.2) that the results are much more consistent for some occupations than for others. For example, the results are more consistent for salesclerks, protective workers, and trades and crafts, than for salesmen and service workers.

Some of these results look very encouraging, but a note of caution must be added. Although not so stated in the report, it is probable that many, if not most of the studies covered by the survey, were studies in which *present employees* had been tested rather than *applicants*. Because of the limitations mentioned before, it is very possible that the validities in some of these cases would be *lowered* if they were established on the basis of *applicant* groups. The results given in

TABLE 6.4

Weighted Mean Validity Coefficients of Personality Inventories for Various Occupational Groups

Mean r	*Total No. of Cases*	*Total No. of* r*'s*	*Occupation Group*
.14	518	8	General Supervisors
.18	6,433	44	Foreman
.25	1,069	22	Clerks
.36	1,120	8	Salesclerks
.36	927	12	Salesmen
.24	536	5	Protective Workers
.16	385	6	Service Workers
.29	511	8	Trades and Crafts

Source: Ghiselli and Barthol, *op. cit.*

Fig. 6.2 and Table 6.4 therefore probably should be considered as over-estimates of the validities that would actually occur in true-life employment situations.

Other Uses of Personality and Interest Tests. This discussion of personality and interest tests deals primarily with their use for employment purposes. It should be kept in mind, however, that such tests have other uses as well. In fact, in many, if not most other situations they would be less likely to have the limitations that have been mentioned above. For example, in vocational guidance or clinical counseling the counselee most typically is genuinely interested in finding out something about himself, and usually he will be relatively honest in responding to the test items. Even in an industrial setting there may be circumstances where the individual being tested would be motivated to be as honest as possible in taking a personality or interest test. This might be the case, for example, with a person who has been hired, but has not been specifically placed on a particular job, or with a person who is being counseled by someone in the company about a personal problem or a job placement problem. In such situations personality and interest tests may be very useful devices.

The point is reiterated here, however, that it is particularly important that such tests be thoroughly validated if they are to be used for employment purposes.

7

ACHIEVEMENT TESTS

Achievement tests, as we have noted, are used to measure the level of proficiency of individuals in some work activity or subject matter. In business and industry such tests are used primarily for the following purposes: (1) the employment of people for jobs for which experienced workers are sought, (2) the transfer and/or promotion of present employees, and (3) in connection with training (as in identifying areas for which additional training is desirable and measuring the results of training that has been given).

Job achievement tests may be of several forms, such as standardized oral questions, written tests with objective questions, or job samples. Some job achievement tests must be given individually; others may be readily administered to groups of individuals. In comparing achievement tests with aptitude tests (as used for employment purposes), achievement tests generally measure how much actual job proficiency the applicant is able to demonstrate at the time he seeks employment, whereas aptitude tests measure the latent capacity to develop proficiency.

Validity of Achievement Tests. The validity of achievement tests is usually established on the basis of *content* validity (in which the appropriateness of the content is determined by "expert judgment"), or on the basis of *concurrent* validity (in which the test is evaluated by showing how well the test scores correspond with measures of some concurrent criterion). The validation on the basis of concurrent validity is by all odds the preferable method, when circumstances make it possible. When this method is used, it is most typically the practice to use two or more job-groups as the criterion groups, such as a group of journeymen in a trade, and a separate group of apprentices in the trade, for purposes of comparison.

ORAL ACHIEVEMENT TESTS

Oral achievement tests are used in some situations, usually in connection with various trades. When used for trades these tests are commonly called *oral trade questions*. Such questions are convenient to administer and simple to interpret. The most extensive research with this type of test has been carried on by Stead, Shartle, and associates. The general procedure followed by this group has been described in detail in their book, *Occupational Counseling Techniques*.[1] Since the many trade questions developed are in use by the various state employment offices, and since the publication of the actual questions used would cause them to lose much of their value, Stead and Shartle publish only a few sample questions to illustrate the method by which the validity of the questions has been determined. This procedure is sufficiently important to the general subject of test validation to warrant a somewhat detailed discussion.

Validation of Oral Trade Questions. The general procedure used by the U.S. Employment Service in validating these questions consisted of finding for each trade a few questions, usually 15, each of which was answered correctly by a large percentage of successful journeymen in that trade, but was answered correctly by a definitely smaller percentage of apprentices in the trade, and by a still smaller percentage of employees in related occupations. When a set of questions for a given trade has been obtained in this manner, it is a simple matter to determine whether an applicant who claims to possess journeyman abilities for that trade actually is able to answer these questions as well as known tradesmen have answered them. The differentiation obtained by such a set of oral trade questions for asbestos workers is shown in Fig. 7.1. The expert asbestos workers vary over a range of 7 to 15 correct answers; apprentices and helpers, of zero to 12; and related workers, of zero to 5. Thus, an applicant who is able to answer correctly only 6 or fewer of the 15 questions is very unlikely to be an expert asbestos worker, regardless of what he may say or feel about his own skill. Even if he answers up to 12 items correctly, he is probably not highly proficient in his trade, because

---

[1] W. H. Stead, C. L. Shartle, *et al.*, *Occupational Counseling Techniques* (New York, American Book Co., 1940).

FORM I

Score	Expert Asbestos Workers (50 Subjects)	Apprentices and Helpers (25 Subjects)	Related Workers (25 Subjects)
15	xxxxxxxx		
14	xxxxxxxxxxxxxxxxx*		
13	xxxxxxxxxxxx		
12	xxxxxx	xx	
11	xx		
10	x	x	
9	x		
8	xx	xx	
7	x	x	
6		xxxxxxxx*	
5		xxx	xxxx
4		xxxx	xx
3		xx	xxx
2		x	xx
1		x	xxxx*
0			xxxxxxxxxx

* Median Score

FIG. 7.1. Distributions of scores for expert asbestos workers, apprentices and helpers, and related workers on a 15-question oral trade test.

the majority of men known to be experts were able to answer 13 or more of the questions correctly.

Factors Affecting the Validity of Trade Questions. The success with which such a set of trade questions will differentiate among applicants possessing different degrees of skill depends upon the very careful selection of questions comprising the test. This selection can be made only by starting with many more questions than will finally be retained and by eliminating every question that does not actually show a differentiation among the several groups tested. Likewise, if the test is to be used on a countrywide basis, the preliminary validation work must be carried on in the several geographical areas where the test is later expected to function. The following question for a roofer illustrates the necessity for considering geographical factors in the validation of the question.

Q. What type of asphalt is used on a flat roof?
A. Flat (F) (low melt).[2]

---

[2] *Ibid.*, 38.

Experts west of the Mississippi River answered this question with unusual consistency, but there tended to be a reduction of consistent answers by experts when the validation took place eastward. Therefore this question could not be used in a test that was intended for countrywide use. In general, questions that could be answered simply, preferably by one or two words, were found to be better than those that called for more extensive answers. Often the best questions were those that could be correctly answered only by terms peculiar to the trade in question, even though such terms might not be found in a dictionary.

Trade questions of this general type have been prepared for most of the standard trades and are in use in many state employment offices.

Although most oral trade tests include questions to which short, concise answers are to be given, true-false questions can also be used. Examples of true-false questions covering the job of heater in a steel mill are given in Table 7.1.

TABLE 7.1

Trade Questions in True-False Form Covering the Job of Heater in a Steel Mill

1. The furnace checkers are always located directly below the car bottom. T F
2. The gas flow and air flow must always come from opposite side of the furnace. T F
3. The only way to tell if the stock gas is burning is to look at the top of stock. T F
4. The purpose of the stock gas is to help the furnace draw. T F
5. In lighting a cold furnace, the rear parts are always the first ones to be opened. T F

Industrial Uses of Oral Achievement Tests. Although oral achievement tests apparently have not been adopted very widely by industry, this testing technique seems to offer considerable promise for measuring achievement. In addition to differentiating different *job* groups (such as journeymen and apprentices), such tests also can differentiate rather well different levels of ability *within* a job. This was indicated, for example, in a study by McCormick and Winstanley[3] of machinists in a shipyard.

Ninety-eight machinists took a 21-item oral test that included questions such as those shown in Table 7.2. The 21 items in this test had been identified from an original test of 95 items as being the ones that were most discriminative. The machinists were divided into three approximately equal groups (A, B, and C) on the basis of ratings by their supervisors on "over-all" job performance. Fig. 7.2 summarizes the results, by indicating the per cent in each criterion group (A, B, and C) whose scores were 13 and above; it shows that the test differentiates quite clearly among the three criterion groups. It was suggested that such a test might well be used for determining the need of individuals for training, for upgrading, for placement or transfer, for selection, for measuring level of ability, and for use in handling grievances that deal with questions of promotions.

---

[3] E. J. McCormick and N. B. Winstanley, "A Fifteen-Minute Oral Trade Test," *Personnel*, September 1950.

TABLE 7.2

Types of Questions Included in Oral Trade Test for Shipyard Machinists

1. *Q.* What material is used in the manufacture of lapping plates?
   A. Cast iron.
2. *Q.* How many turns of a handle are required to make a half turn of the most common index head?
   A. 20.
3. *Q.* What is the meaning of the Brinnel test?
   A. A method of determining the hardness of metal.
4. *Q.* For what purpose would you use transfer calipers?
   A. Used to determine size of recesses and places where the legs of the calipers must be moved to get them out.

Source: McCormick and Winstanley, *op. cit.*

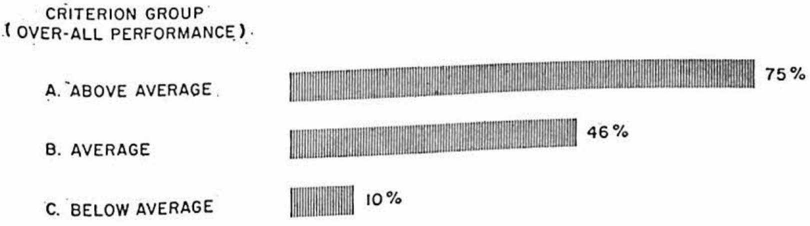


FIG. 7.2. Per cent of machinists in each of three criterion groups scoring at or above critical score of 13 on oral trade test. (Adapted from McCormick and Winstanley, op. cit.)

JOB SAMPLE TESTS

The job sample (also called job miniature or work sample) is an achievement test that consists of trying out the individual in a test situation that reproduces all, or an important sampling of, the actual operations that the job itself requires. Such a test usually utilizes equipment such as that used on the job. If the equipment actually used on the job is hazardous, simulated equipment is sometimes constructed which eliminates the hazards of the usual machines. A scoring procedure is developed, and norms of experienced and inexperienced workers are usually obtained in the test situation as a basis of evaluating scores of persons taking the test, such as applicants.

A Miniature Punch Press. An example of this approach is the miniature punch press described by Tiffin and Greenly.[4] This test is used in the selection of punch press operators. The apparatus, which is illustrated in Fig. 7.3, is a replica in all essential features of a small industrial punch press. It differs from a real press in that the punch is located in a vertical bearing and is held down only by a spring. This feature prevents the punch itself from descending when an obstacle is encountered. When this occurs, the punch remains stationary while a mechanical counter records an error or mispunch. The test is administered by having an applicant put through the press 200 prepunched pieces of galvanized sheet iron. The time required to feed these pieces is recorded by means of a stop watch. During the test period the mechanical counter records the number of errors or mispunches. The test thus results in simultaneous time and error scores.

The validity of this test has been studied by comparing the average performance of different groups of persons to whom the test had been given. The curves in Fig. 7.4 summarize the results obtained for three specific groups. These curves show the relation between mispunches or errors in punching 200 plates and the time in minutes required to punch the 200 plates. As might be expected, the errors decrease as the time increases for all three groups of persons upon whom these results are based. It is interesting to note, however, that the curve for the students and the one for the insulation stripping machine operators are almost identical mathematically and simply represent different segments of the same curve. This

---

[4] J. Tiffin and R. J. Greenly, "Experiments in the Operation of a Punch Press," *Journal of Applied Psychology*, 1939, 23, 450–460.

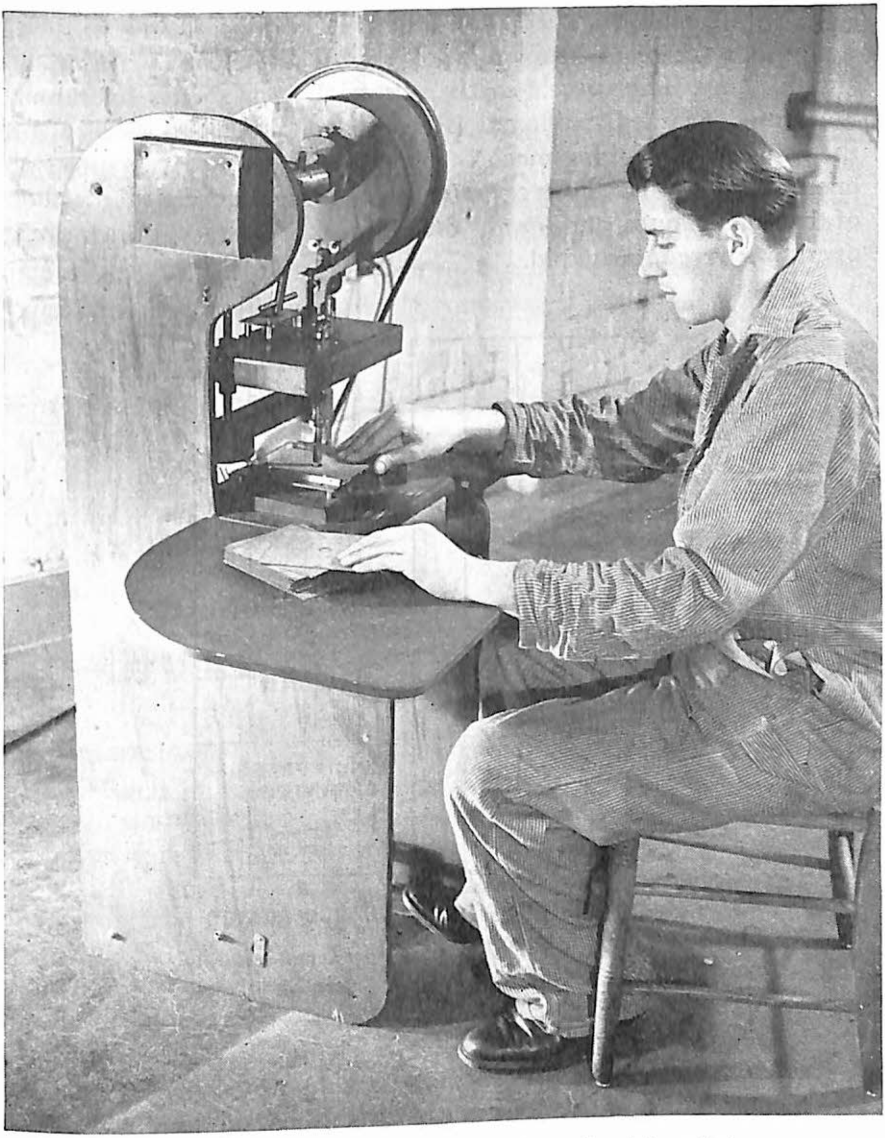


FIG. 7.3. Applicant being given the Miniature Punch Press Test.

suggests that the students, though punching much more rapidly and hence making many more errors, are not significantly different from the insulation stripping machine operators in genuine ability. The employees in this latter group, though offering no experience in the operation of a punch press, were industrial employees and, as such, were more accustomed to the need for careful, slow operation of any machine. The curve for the industrial punch press operators, however, is markedly different from that of either the students or the insulation stripping machine operators. For any given speed of operation the punch press operators were more accurate than either of the other groups, and for any given level of accuracy, punch press operators were more rapid.

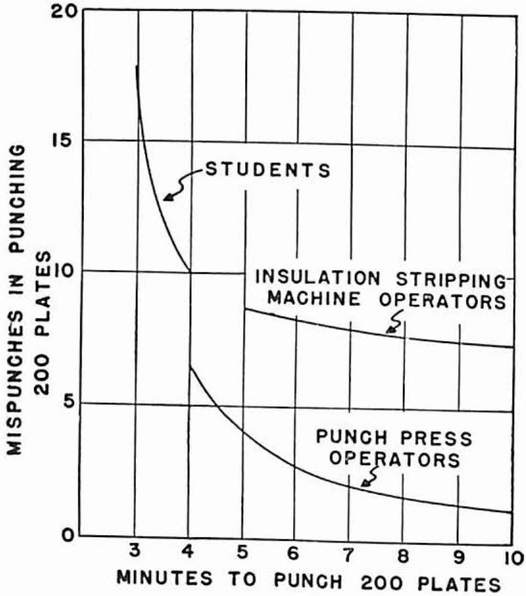


FIG. 7.4. Relation between speed and accuracy for three groups of subjects in operating a small punch press.

From a knowledge of the speed and accuracy of a given applicant and a comparison of this information with the data graphed in Fig. 7.4, it is possible to determine the status of the applicant in comparison with the corresponding test performance of persons who are known to be experts on this job.

The method illustrated by the miniature punch press has been utilized in selecting employees for numerous other jobs. For example, in hiring persons for such jobs as packaging, inspecting, and certain types of machine operation, it is often extremely helpful for the personnel manager to obtain a sample of the quality and speed of work that the applicant is able to perform.

A Performance Test for Fork-Lift Operators. An example of a very different type of job sample test is one that has been developed by the Philadelphia Quartermaster Depot for fork-lift operators.[5] A fork-lift truck is used for moving materials in a warehouse or around an industrial plant. The test consists of driving a loaded fork-lift truck around a standard driving course as shown in Fig. 7.5. The "walls" and the obstacles are constructed from pallets 32 by 40 inches, painted yellow to contrast with the floor. Without going into detail, it should be pointed out that the equipment used and instructions should be standardized. The person giving the test has a check sheet on which he records his observations. This check list includes a listing of 42 poor operating techniques such as "Did not start in low gear," "Started jerkily," "Scraped side walls of tires," "Lowered load too quickly," and "Number of pallets displaced." Scoring is based on the number of errors made, as recorded on the check list.

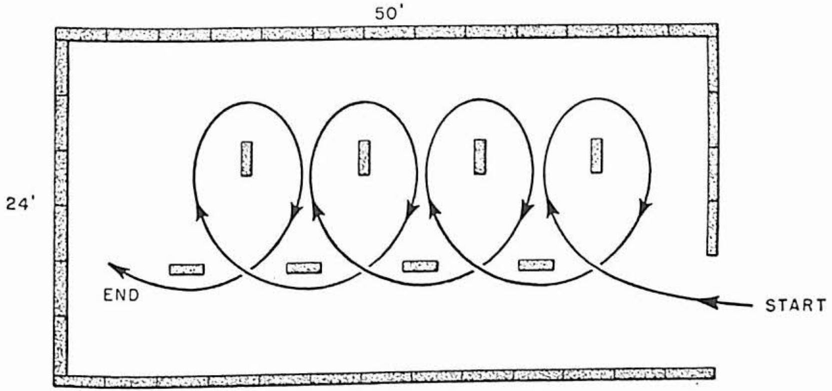


FIG. 7.5. Standard driving course diagram for performance test for fork-lift operator. (Test prepared by U.S. War Dept., Office Quartermaster General, Philadelphia Quartermaster Depot, Testing Station.

Although it might appear that many jobs would not lend themselves to job sample tests, it can be seen from an example such as this that the use of imagination and inventiveness can result in the development of tests that can be objectively scored for a wide variety of jobs.

---

[5] This test was prepared by the U. S. War Department, Office Quartermaster General, Philadelphia Quartermaster Depot, Testing Section.

The Job Sample in Clerical Work. Another area in which the job sample method of testing has resulted in very satisfactory results is the selection of employees for stenographic, clerical, and secretarial positions. One of the most widely used tests of this type is the series known as the Blackstone Stenographic Proficiency Tests.[6] This series includes a stenography test and a typewriting test. The stenography test is made up of seven parts, namely, knowledge of English grammar (which includes punctuation, capitalization, and spelling); syllabification; office practice; alphabetizing; abbreviations; knowledge of business organization; and ability to take dictation and shorthand and to make the correct transcription. Separate norms are available for each of these seven parts so that in hiring an applicant for a given job it is possible to match the test results with a job analysis of the work to be performed. The second part of the Blackstone series consists of a typewriting test. This part requires the applicant to copy a standard page of typewritten material. The result may be scored separately for speed and accuracy and the norms available permit a rapid determination of the exact level of proficiency possessed by an applicant.

The Purdue Clerical Adaptability Test[7] has been discussed in Chapter 5 in connection with aptitude tests. In many circumstances, however, certain parts of this test can be used effectively as achievement tests.

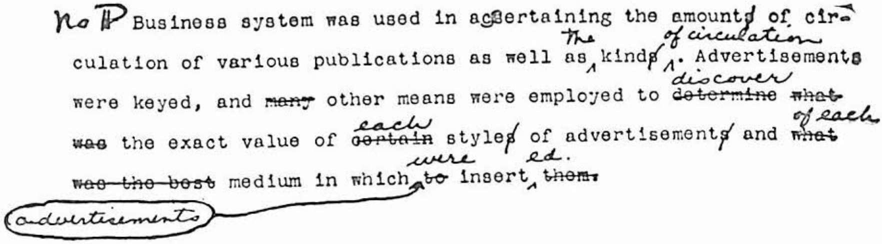
A standard test used in measuring proficiency in the use of a typewriter is the Thurstone Examination in Typing.[8] A part of this test consists of a page of corrected copy that the applicant is asked to type, making the indicated corrections. A sample of the material to be copied is shown in Fig. 7.6. The ability to copy corrected material of this kind, though frequently called for in many stenographic positions, seems to be quite different from the ability to make an exact copy of material requiring no corrections. For example, in one em-

---

[6] *Blackstone Stenographic Proficiency Tests* (Yonkers, N.Y., World Book Company, 1932).

[7] Distributed by the University Book Store.

[8] *Thurstone Employment Tests, Examination in Typing* (Yonkers, N.Y., World Book Company, 1922).



no ¶ Business system was used in ascertaining the amounts of cir-
the of circulation
culation of various publications as well as kinds . Advertisements
discover
were keyed, and ~~many~~ other means were employed to ~~determine~~ ~~what~~
each of each
~~was~~ the exact value of ~~certain~~ styles of advertisements and ~~what~~
were ed.
~~was the best~~ medium in which ~~to~~ insert ~~them.~~
advertisements

FIG. 7.6. Part of Test I of the Thurstone Typing Test. The person tested is required to copy this material, making the indicated corrections.

ployment office a girl who had recently won a state contest for speed in typing was unable, when tested with the Thurstone test, to obtain even an average score in typing this corrected copy. Apparently her ability consisted of a strictly mechanical and automatic method of "copying the copy." Such ability might be very useful in winning state contests, but it is of little use in a business situation because one is seldom asked to copy material that is perfect at the outset. A test of the Thurstone type, therefore, is a much better measure of the kind of ability called for in a business or industrial job than is a test that requires only straight copying of material.

An excellent worksample test of transcription and typing has been recorded on standard phonograph records by Seashore and Bennett.[9] Letters are dictated from records at varying speeds. The applicant takes the material in shorthand and later types the letters in standard form. The score on the test is determined by the appearance and accuracy of the final typed letters.

WRITTEN ACHIEVEMENT TESTS

Although many achievement tests are "tailor-made" for particular uses by a company or other organization, there are some that have been developed and standardized for widespread use. This is particularly the case with written achievement tests.

The Purdue Vocational Tests. The Purdue Vocational Tests[10] comprise a series of this type. These tests, although developed primarily for use by vocational teachers in public school systems, are being used also by industrial personnel men for the purpose of determining the qualifications of applicants for jobs connected with the various trades. Several standardized tests are now available in this series.

---

[9] H. G. Seashore and G. K. Bennett, *The Seashore-Bennett Stenographic Proficiency Test* (Psychological Corporation, 1946).

[10] Distributed by Science Research Associates, Chicago, Illinois.

Technical Information in Machine Shop.[11] This test consists of 132 questions (partly multiple choice and partly matching) that cover material well known to any qualified machinist and that should be reasonably familiar to any boy who has received adequate vocational instruction in this field. Four examples of the multiple-choice items follow:

In each of the multiple-choice statements listed below there are four possible answers, but only one is correct. Read each statement carefully before making your choice of answers.

By tolerance in machine work is meant (*a*) the allowance, as for the oil film of a bearing, (*b*) the amount of variation either above or below a certain basic measurement that will be acceptable, (*c*) the amount of stock left for polishing, (*d*) the amount of stock left for grinding.. (   )

If it fits, the wrench which will probably do the least harm to the corners of a nut is (*a*) an adjustible, (*b*) an alligator, (*c*) an open end, (*d*) a socket wrench.................................. (   )

One of the best bearing metals contains antimony, tin, and copper. This metal is called (*a*) bronze, (*b*) brass, (*c*) babbitt, (*d*) lead.. (   )

Standard taper sizes are designated by (*a*) decimals, (*b*) fractions (*c*) numbers, (*d*) letters.................................. (   )

The test covers such topics as hand tools, bench tools, bench work, lathe, milling machine, shaper, planer, and drill press. The test is arranged so that subscores may be readily obtained for bench operations and for the operation of the grinder, planer, lathe, and milling machine. On the basis of odd-*vs.*-even-item correlations, the reliability of the various scores was as follows.[12]

Bench operations	.87
Grinder	.90
Planer	.89
Lathe	.80
Milling Machine	.85
Total Score	.96

---

[11] H. F. Owen, C. C. Stevason, H. G. McComb, and C. D. Hume, *Technical Information Test for Machinists and Machine Operators*, Form BH (Chicago, Science Research Associates, 1949).

[12] *Preliminary Manual for the Purdue Test for Machinists and Machine Operators* (Chicago, Science Research Associates).

The published norms were obtained on the basis of a 50-minute time limit. The scoring is simple and rapid. Separate norms are available for students or applicants offering different amounts of vocational training or practical experience.

Blueprint reading test.[13] This test consists of 79 questions covering basic principles involved in reading and interpreting blueprints.

Technical information in electricity.[14] Another test in the Purdue Vocational Series deals with industrial electricity. This test is available in two forms, A and B, each consisting of 65 multiple-choice and matching type questions covering material of a practical nature in the field of electricity. Four sample questions from this test follow:

In the following section of this test select the word or phrase after each incomplete sentence which makes the most nearly true statement. Record the number of the choice made.

A bathroom fixture should never use a brass shell socket because it may (1) short circuit and blow fuses, (2) draw excessively high currents, (3) give one a dangerous shock, (4) become corroded and open circuit. . . . . ( )

Conductors of electricity are most often made of (1) iron, (2) brass, (3) lead, (4) copper. . . . . ( )

An electric bell depends for its operation upon (1) the heating effect of electricity, (2) the mechanical effect of electricity, (3) the magnetic effect of current flow, (4) the impedance of the circuit. . . . ( )

An accidental connection between two wires of opposite polarity is called (1) a short circuit, (2) a parallel circuit, (3) a series circuit, (4) a ground. . . . . ( )

This test covers such topics as common electrical circuits, measuring units, conductors, and common electrical devices. Its reliability, computed by correlating scores on Form A with scores on Form B, is .91. The time limit is 25 minutes.

Technical information in industrial mathematics.[15] In its construction, scoring, and administration, this test is similar to the tests just described. It deals primarily with operations of an arithmetical, or simple mathematical, type that a tradesman in industry is likely to encounter. The test is available in two matched forms, A and B. A preliminary computation, based on a group with a rather narrow range of scores, gave a correlation of .82 between the two forms. A sample from this test reads as follows:

---

[13] H. G. Owen and J. N. Arnold, *Purdue Blueprint Reading Test* (Chicago, Science Research Associates).

[14] C. W. Caldwell, H. R. Goppert, H. G. McComb, and W. B. Hill, *Technical Information Test for Electricians* (Chicago, Science Research Associates).

[15] C. H. Lawshe, Jr. and Dennis H. Price, *Purdue Industrial Mathematics Test* (University Book Store, West Lafayette, Indiana).

> This test consists of mathematical problems that are found in various kinds of industrial work. Each problem has FOUR suggested answers, but only ONE is correct. Write an X over the correct answer.
>
> 1. The difference between $\frac{1}{3}$ and $\frac{1}{4}$ inch is: .... (1) (2) (3) (4)
>   (1) $\frac{1}{6}''$ (2) $\frac{1}{8}''$ (3) $\frac{1}{10}''$ (4) $\frac{1}{12}''$
> 2. If 28 man-hours are required to build 4 tables, how many man-hours will be needed to build 5 tables? .... (1) (2) (3) (4)
>   (1) 32 (2) 35 (3) 38 (4) 41
> 3. If a piston has an area of 6 square inches, and a force of 30 pounds is required to move the piston, the pressure per square inch is: ...... (1) (2) (3) (4)
>   (1) 5 (2) 6 (3) 7 (4) 8
> 4. If $3a + 7 = 22$, then "a" equals: .... (1) (2) (3) (4)
>   (1) 3 (2) 4 (3) 5 (4) 6

Engine lathe operation. The Purdue Trade Information Test in Engine Lathe Operation[16] consists of 74 objectively scored questions dealing with the operation of this machine tool.

Welding. The Purdue Trade Information Test in Welding[17] consists of 70 objectively scored questions covering both gas and electric arc welding.

Carpentry. The Purdue Trade Information Test in Carpentry[18] consists of 90 objectively scored questions in the field of carpentry.

Interview Aids. Another series of simplified achievement tests has been published under the series title *Interview Aids.*[19] These include "Can You Read a Micrometer?" "Can You Read a Scale?" and "Can You Read a Working Drawing?"

Achievement tests of this sort in the different trade areas enable an employment manager to obtain a thorough picture of the strong and weak areas of information possessed by applicants. Because the tests may be given to many applicants simultaneously, they can be made much longer than oral trade tests and therefore are much more exhaustive than the latter in covering a technical field.

---

[16] R. Cochran and J. Tiffin, *The Purdue Trade Information Test in Engine Lathe Operation* (University Book Store, West Lafayette, Indiana).

[17] J. Tiffin and W. B. Griffin, *The Purdue Trade Information Test in Welding* (University Book Store, West Lafayette, Indiana).

[18] J. Tiffin and R. F. Mengelkoch, *The Purdue Trade Information Test in Carpentry* (University Book Store, West Lafayette, Indiana).

[19] Distributed by Science Research Associates.

USES OF ACHIEVEMENT TESTS

Some of the uses of achievement tests have been mentioned or illustrated briefly above. A few additional illustrations will be given here.

Employment. One of the most important uses for achievement tests is in the employment of applicants. A typical employment problem for many companies is that of selecting apprentices for the various skilled trades. This is an important problem because of the time and expense involved in apprentice training. Since many applicants for apprenticeships have had high school courses in areas related to the trade involved, achievement tests have been found to pay substantial dividends in the selection of men for the apprenticeships. In such a case an achievement test covering a certain subject area is useful in the prediction of continued development in the subject area in question. (In one sense, such tests are then used more as aptitude tests than as achievement tests.)

The desirability of using a test in such a circumstance is illustrated in Fig. 2.2 in Chapter 2 on individual differences. This shows that among 112 applicants for machine shop apprentice training a considerable number made a score of less than 40 (number of items correct) and a scattered few made a score of less than 20 on the machine shop achievement test. Although most of these applicants had had at least one semester of machine shop instruction, and many had had much more than this amount of training, 25 per cent of them did so poorly on the machine shop achievement test that they were below the lowest 10 per cent of students in vocational classes. It was clear, therefore, that among these applicants were a considerable number who would be very unlikely, even with prolonged training, to become expert machinists. The company concerned set 90 items correct as the critical score for hiring for this job and was able, by this process, to select for apprenticeships boys who made very rapid strides in developing the necessary skill to become expert machinists.

Transfer and Promotion of Employees. The problem of transfer and promotion is one that continually confronts every operating management. The use of appropriate achievement tests can aid in identifying individuals who have the skills or knowledges required for other jobs. Note that we mentioned *appropriate* achievement tests. The fact of a person being proficient on his *present* job often gives no assurance of abilities for *other* jobs. In other words, an achievement test used for transfer or placement purposes must be valid *for the job for which the individual is being considered.*

The use of tests for transfer and promotion purposes should supplement, rather than replace, other considerations such as seniority, being "in line" for the job, and the considered judgments of management representatives regarding the qualifications of the individual in question.

Training. Many companies are devoting more and more time to the systematic training of both new and old employees. The need for such training is a natural result of the continuous technological

TABLE 7.3

Part of an Achievement Test Covering Information That Should Be Known to Employees Seeking Transfer or Promotion

Instructions: On the right is a list of colors. On the left is a list of the materials carried in pipes in this mill. You are to show how well you know the color code by matching each color with the figure or figures which you find before the appropriate materials. Mark your choice in the parentheses at the extreme right. The first one is correctly marked to show you how it should be done. There will be some colors unused. Some others will be used twice.

Materials	*Colors*	
1. Stabilized Gas	Aluminum	(6) ( )
2. Steam	Black	( ) ( )
3. Hot Water	Brown	( ) ( )
4. Cold Water	Bright Red	( ) ( )
5. Coke Oven Gas	Ceiling Blue	( ) ( )
6. De-oxidized Gas	Dark Purple	( ) ( )
7. Compressed Air	Dark Red	( ) ( )
8. Natural Gas	Gray	( ) ( )
9. Fuel Oil	Green	( ) ( )
10. Sulphuric Acid	Blue	( ) ( )
11. Farval Grease	Pink	( ) ( )
12. Bowser Oil	Olive Green	( ) ( )
13. Gasoline	Orange	( ) ( )
14. Kemp Lines	Lavender	( ) ( )
15. Water Fire Lines.............	White	( ) ( )
16. Foamite	Yellow	( ) ( )
17. Hydraulic Lines		

changes occurring in modern industry. No matter how well qualified an employee may be today, technological change in methods or processes may require that he be completely retrained tomorrow. The systematic use of technical information tests among present employees furnishes a convenient means of determining those areas in which training is needed. An example of this use of information tests is shown in Table 7.3. This table consists of a set of matching items to determine whether the employees are familiar with the color code used in the plant in which they work. Several other important areas were covered in the tests. The content of the training program that followed the administration of this test was based largely upon the results obtained.

An interesting use of achievement tests in training is described by Taylor,[20] the report dealing with an automotive mechanics course at an Army Ordnance School. The Army personnel going through this training course typically went through three four-week phases of training, adding to a total of 12 weeks. The three phases increased in complexity.

A series of fifteen tests was developed for use in the training program, including both written and job sample tests. An analysis of experimental results with these resulted in the selection of five of them for the purposes to be described. These five included an Auto Mechanics Multiple Choice Test, a Tool Usage Film Strip Test, an Engine Ignition Checks Test, a Hand Tools Checks Test, and an Auto Mechanic Experience Check List. From a total group of 807 incoming trainees, the 178 with the highest scores on these tests were identified. These students "skipped" the first phase and entered training at the beginning of the second phase. At the end of the training (in this case 8 weeks), their performance was compared with that of the remaining 629 students who had gone through the full 12 weeks. Performance at the end of training was compared on the basis of final grades in the training course. This comparison, shown in Fig. 7.7, indicates very clearly that those who skipped the first four weeks were, as a group, superior to the others at the end of training, despite the fact that they had had only 8 weeks of training. For this particular group of 178 men, 4,984 man days (13.7 man years) were saved as a result of eliminating the first four weeks of

---

[20] C. W. Taylor, "Pre-testing Saves Training Costs," *Personnel Psychology*, 1952, 5, 213–239.

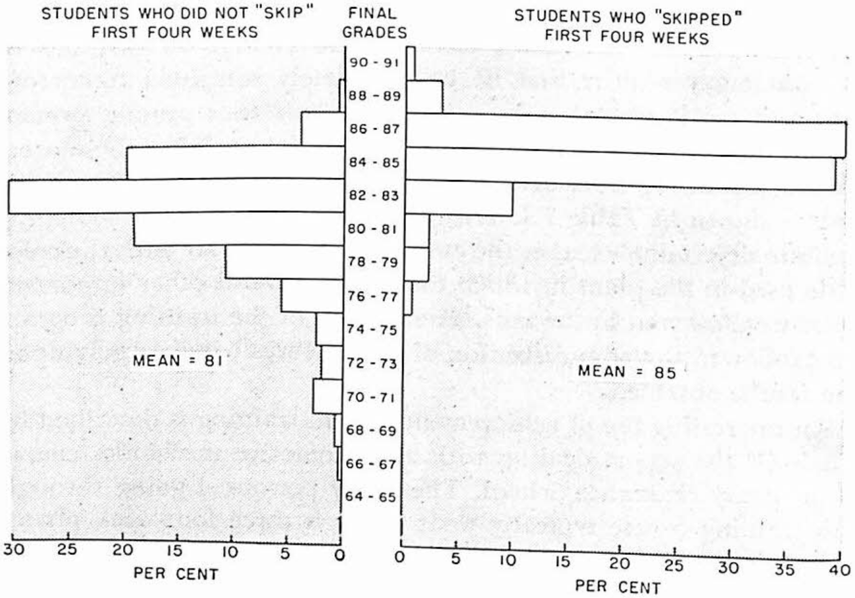


FIG. 7.7. Distributions of final average scores for automotive mechanic students who "skipped" the first four weeks of training and for those who did not. Those who "skipped" were individuals who scored highest on five achievement tests taken at the beginning of training. (From Taylor, op. cit.)

training for them. This reduction in training was made possible, of course, by using the achievement tests for identifying the individuals whose initial level of achievement was highest.

Measurement of Vocational Achievement. Although our primary interest in achievement tests is concerned with their use in business and industry, it should be pointed out that such tests also have great utility in vocational training work. Every vocational teacher expects to place his students in industrial jobs. The success of these students in such jobs, however, depends largely upon the adequacy of training they have received while students. Teachers in the field of general education have long made use of standardized achievement tests to determine student achievement in the several school subjects and to compare the achievement of students in various school systems, under different types of instruction, and in different geographical localities. Administrators and teachers of vocational subjects are now beginning to make similar use of objective achievement tests, and there is every reason to believe that the judicious use of such tests in this area will be fully as valuable as in the field of general education. The several trade achievement tests discussed in this chapter are ideally suited to this purpose. Since vocational demands differ from one industry to another and from one industrial center to another, it is often wise to build tests that are "tailor-made" for the particular situation. Such tests, when constructed through the joint efforts of vocational teachers, school administrators, vocational co-ordinators, and representatives of the industries concerned, furnish an ideal means of facilitating the co-ordination between industries and schools that every community desires to encourage.

8

INDUSTRIAL MERIT RATING

Merit rating is a systematic evaluation of an employee by his supervisor or by some other qualified person who is familiar with the employee's performance on the job. Merit ratings are usually made by means of a standardized form that is adapted to the needs of the particular industry. Usually the ratings are made at periodic intervals. A merit rating thus becomes a permanent part of an employee's record with a given company, and, at least in theory, is a part of the record that may be used by management in subsequent promotion, transfer, or layoff.

THE GROWTH OF MERIT RATINGS

A survey of 400 companies carried out by the National Industrial Conference Board[1] indicated that about one half had employee merit rating plans. Although this is not a majority of the companies surveyed, it represents a significant proportion. Despite certain criticisms that have been made regarding merit rating systems, it seems probable that merit rating will remain an important part of personnel administration programs of many companies. Such an assumption justifies a rather careful discussion of this topic.

One point should be made clear immediately: although the term "merit rating" is new, the rating of men by supervisors is by no means a new development. Supervisors have *always* rated men, and it is no doubt true that the ratings made in a random, slipshod, and unsystematic fashion, unrecorded and undefended, whether valid or not, have in the past been just as important in determining whether a given employee should hold his job as any rating made by means of a modern merit rating chart. The changes that came with systemmatic merit rating, then, did not involve making ratings where none had existed before; rather they involved the transfer of ratings from haphazard, random, and frequently irresponsible judgments of supervisors made perhaps during the heat of a quarrel, to ratings made calmly, deliberately, systematically, and in a manner that made the ratings, if not completely comparable from one employee to another, at least much more comparable than were the older, haphazard evaluations of employees by supervisors. The question, then, is not whether supervisors should rate their employees—this always has been and probably always will be done—but whether the use of a formal merit rating system will increase the value of such ratings both to management and to employees.

---

[1] "Personnel Practices in Factory and Office," *Studies in Personnel Policy*, No. 145, National Industrial Conference Board, 1954.

PURPOSES OF MERIT RATING

Various reasons have been given by company executives for using merit rating systems. In a survey reported in *The Management Record*,[2] fifteen different uses of merit ratings were listed. In general, however, these purposes seem to fall into two major categories: "administrative" and "self-improvement" (or clinical). In the second category, the emphasis is upon helping employees to understand their strengths and shortcomings, so that they can have some basis for self-improvement.

A summary of the uses of merit ratings suggests that the following are the most important.

Administrative Uses of Merit Ratings.

Promotion. It is probable that this is the most important administrative use of merit ratings. It is to the common interests of both management and employees to promote employees into positions where they can most effectively utilize their abilities. It is a disservice to both parties to promote employees into positions where they cannot perform effectively at the time in question. A properly developed and administered merit rating system can aid in determining whether individuals should be considered for promotions. There is, however, an important distinction to be made in using merit ratings for this purpose. Such ratings should differentiate between the *performance* of the individual *on his present job,* and his *potentiality* for performance on a *higher-level job.* The ability to perform effectively on one job does not necessarily give assurance of an employee's potentialities for greater responsibilities. These potentialities need to be evaluated separately.

---

[2] Summarized in "Appraisal of Job Performance," *Studies in Personnel Policy*, No. 211, National Industrial Conference Board, 1951.

Other personnel actions. There are occasions in most companies when it is necessary to consider various types of personnel actions such as *transfers, layoffs,* and *discharges.* In some cases such actions may be necessary because of unsatisfactory performance on a present job. In other cases the actions may be necessary because of economic conditions over which the company has no control, or because of changes in the production processes.

Although most companies like to avoid taking such actions as layoffs, discharges, and transfers necessitated by inadequate performance, when they become necessary, such actions must be faced squarely. If the employees in question, as well as others, realize that fairness and consideration serve as the keynote of personnel policy, such actions may cause less criticism than if they are based on whim or expediency. When properly used, merit ratings can serve as the basis for fairness and consideration in taking such actions.

Wage and salary administration. In some companies, merit ratings are used as a basis for granting increases in wages or salaries. In some cases both merit and seniority are used in combination in granting wage or salary increases.

Training. An appropriate merit rating system can be useful for training purposes in two ways. In the first place, it can aid in identifying areas of skills or knowledges in which numerous employees are not up to par, thus pointing up general training deficiencies which presumably should be corrected by additional training. In the second place, ratings can help to identify individuals who may need additional, special training.

Criteria for personnel research. As indicated in Chapter 2, merit ratings frequently are used as criteria for personnel research purposes such as in test validation.

Employee Self-Improvement. For a merit rating system to serve this purpose, it is necessary that each employee know where he stands with his supervisor. This involves an interview between a supervisor or some other company representative and the employee in question. It is very difficult for some supervisors to bring themselves to talk with their employees, especially with employees whose work is unsatisfactory. An improperly conducted interview can do more harm than good. The objectives of employee self-improvement are laudable, but a company needs to consider carefully the desirability of inaugurating a program that will fail or succeed depending upon the abilities of supervisors (or others) to conduct interviews with the employees in question.

A very strong case can be made out to provide for *someone* to discuss the ratings of employees with them, in order, as Davis[3] has pointed out, to avoid misinterpretation. Moore[4] has remarked that one of the most valuable uses of a rating program is to identify the weaknesses of employees that may be corrected. Knowles[5] also has emphasized that employees should be informed of their ratings and encouraged to overcome their shortcomings. Armstrong[6] has outlined several principles that may be advantageously followed in discussing ratings with employees, namely:

> 1. Criticism and reprimand should be based upon facts and should avoid personalities.
> 2. Public criticism is rarely justified.
> 3. Private reprimands should be preceded by a statement of some of the more desirable traits of the employee.

The experience of many industries has been that employees should be informed of their ratings by some production man—either the foreman himself or the departmental head. Production men, however, often require special training before they are able to do this smoothly and effectively. Many supervisors who are good production men are not very skillful in discussing a man's weak points face to face with him. To do so successfully calls for tact, objectivity, and a sincere interest in helping the man as well as the business. Many companies have found it wise to devote an extended series of training conferences to teaching supervisors how to inform their men of weak spots without offending the men. When this result has been accomplished and when a supervisor is able to talk face to face with his men about their weak as well as their strong points, a long step has been taken toward solidarity in the working group and toward the upgrading of men who otherwise might, for lack of information, either remain exactly where they are or actually regress.

---

[3] H. A. Davis, "Inefficient Efficiency Rating," *Personnel Journal*, 1944, **22**, 268–270.

[4] H. Moore, "Real Use for Rating Scales," *Personnel Journal*, 1942, **21**, 165–170.

[5] A. S. Knowles, "Merit Rating and Labor Management," *Personnel*, 1940, **17**, 29–42.

[6] T. O. Armstrong, "Talking Your Ratings," *Personnel*, 1943, **20**, 112–115.

If an employee, either for lack of experience or lack of information, continually performs his job in the wrong way, the supervisor will be doing both the employee and the company, as well as himself, a favor by bringing this matter to the employee's attention in the proper manner. Unless a supervisor is required periodically to evaluate all of his employees with regard to their ability in various directions, it is quite likely that many aspects of an employee's performance that could easily be improved, and that the employee himself would like to improve if he were made aware of his shortcomings, will go on from day to day without correction on his part. To inform an employee of his strong and weak points is not only a reasonable and fair thing to do but is also good business for any management that is sincerely interested in having its employees perform their jobs in the best possible way.

CONSIDERATIONS OF MERIT *VS.* SENIORITY

Although the attitudes of unions toward merit ratings vary a great deal, by and large they tend to take a dim view of merit ratings when used for administrative purposes such as promotions, layoffs, and wage increases. Historically, unions place primary emphasis on seniority as the basis for such actions. Seniority, of course, can be objectively determined, whereas merit ratings, by their nature, are based on subjective judgments.

The provisions of union contracts regarding the use of seniority and merit ratings for personnel actions vary. Although some such contracts provide that most personnel actions will be based exclusively on seniority, most of them include provisions such as the following: "Promotions within the bargaining unit will be made on the basis of merit, other things being equal,"[7] or that seniority shall govern promotions when the "ability, skills, and job performance of applicants are equal." Usually the burden of proof for characterizing relative merit of employees in such cases lies with management.

---

[7] "Appraisal of Job Performance," *Studies in Personnel Policy*, No. 121, National Industrial Conference Board, 1951.

The extent to which seniority is a factor in various types of personnel actions in non-union companies is summarized in Table 8.1, which is based on a survey among 130 companies.[8] It can be seen that even non-union companies place a great deal of emphasis on seniority in taking various personnel actions. However, of a total of 110 of these companies (with less than 400 employees each) only four used seniority as a *sole* basis for granting promotions. More typically, various personnel actions were based in part on other considerations. Eighty-two per cent of the companies, for example, considered merit or ability in conjunction with seniority. Other factors that were taken into account by some companies included marital or family status, physical fitness, hardship, attendance, and cooperativeness.

TABLE 8.1

The Seniority Factor in Different Situations

Situation where seniority is a factor	*Companies*	
	No.	*Per cent*
Layoffs	116	89
Promotion	107	82
Rehiring after layoff	96	74
Transfer	69	53
Others	7	5

Source: Adapted from "Seniority Systems in Non-union Companies," *op. cit.*

MERIT RATING SYSTEMS

The individual merit rating plans of specific companies vary a great deal, but they usually fall into one or another of certain basic types. Some plans embody certain aspects of two or more basic methods. The most important basic systems are listed below:

1. Rating scales.
2. Employee comparison systems.
    a) Rank-order system.
    b) Paired comparison system.
    c) Forced distribution system.

---

[8] "Seniority Systems in Non-union Companies," *Studies in Personnel Policy*, No. 110, National Industrial Conference Board, 1950.

3. Check lists.
    a) Weighted check list.
    b) Forced-choice check list.
4. Critical incident technique.

The four categories above are all based on essentially different rating procedures. Since each of the various methods has certain inherent advantages and disadvantages, there is no single "best" method; rather, one method may best serve one purpose, and another method may best serve a different purpose.

Rating Scales. Rating scales, also referred to as "the chart system," are the most widely used type of merit rating system. The basic principle of this method provides for the rating of employees on each of a number of different traits or worker characteristics. There are two primary variations in the manner in which the ratings may be made. In the *graphic* rating scale, a line represents the range of the trait or characteristic, and the rater places a check mark at the position along the scale that he considers to represent the degree of the trait for the employee being rated. An example is shown in Fig. 8.1.

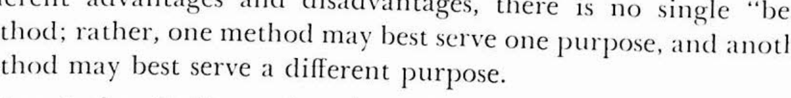


FIG. 8.1. Illustration of graphic rating scale. The rater places a check mark at the position on the scale that represents his judgment of the degree of the trait possessed by the employee being rated.

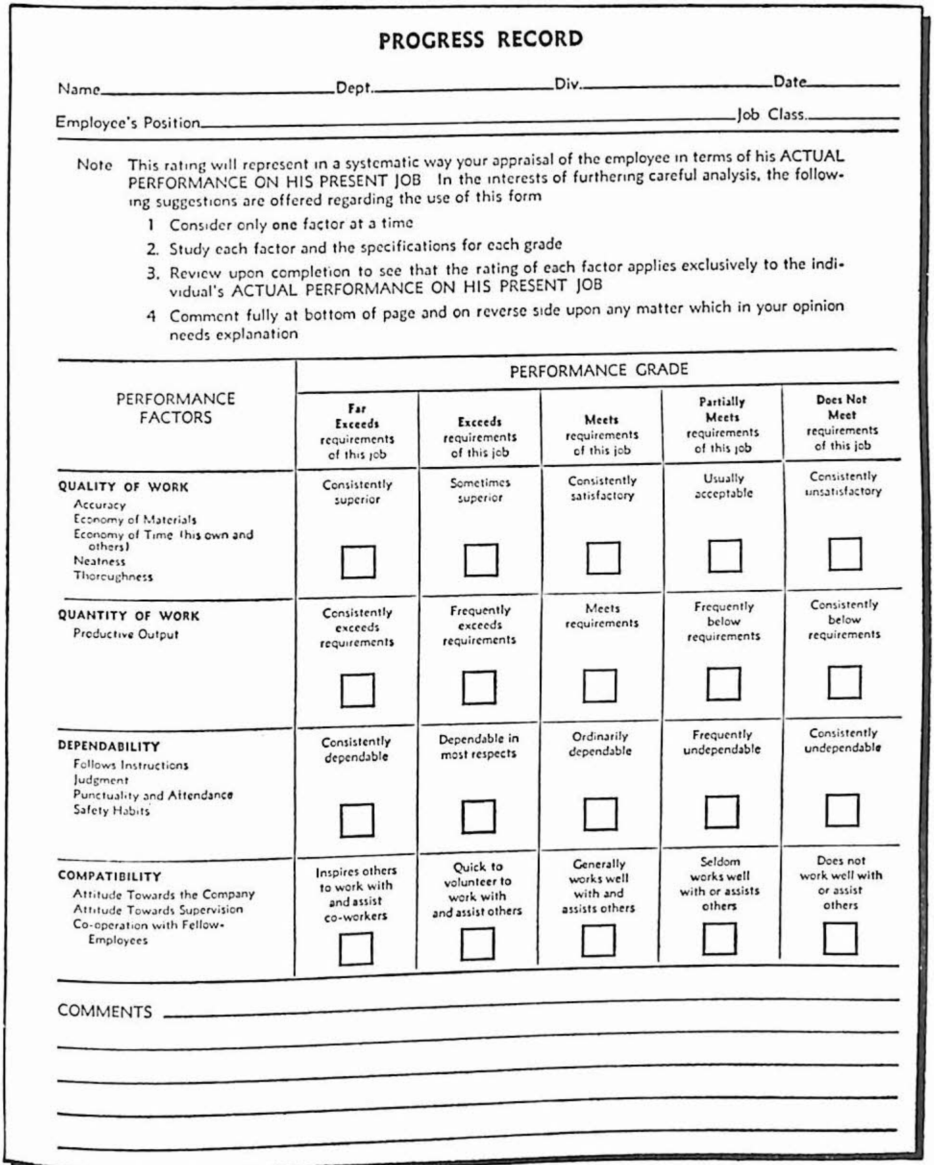
The *multiple-step* rating scale provides for rating on each trait in any one of a number of "degree" categories, as shown in Fig. 8.2.

The rating systems used in various companies differ widely in the number of traits to be rated and in the particular traits used.[9] The variety of traits may be seen from an analysis of the merit-rating charts of eighteen companies[10] as shown in Table 8.2. The companies are referred to anonymously by number across the top row of the

---

[9] A summary of the steps that should be taken in developing a merit-rating system has been published by J. E. Zerga. See "Developing an Industrial Merit-Rating Scale," *Journal of Applied Psychology*, 1943, **27**, 190–195.

[10] Starr and Greenly, *op. cit.* This article refers to a survey of 16 companies, but the tabulated results show an analysis of 18 merit-rating charts.



PROGRESS RECORD

Name__________ Dept.__________ Div.__________ Date__________

Employee's Position__________ Job Class.__________

Note This rating will represent in a systematic way your appraisal of the employee in terms of his ACTUAL PERFORMANCE ON HIS PRESENT JOB In the interests of furthering careful analysis, the following suggestions are offered regarding the use of this form

1 Consider only **one** factor at a time
2. Study each factor and the specifications for each grade
3. Review upon completion to see that the rating of each factor applies exclusively to the individual's ACTUAL PERFORMANCE ON HIS PRESENT JOB
4 Comment fully at bottom of page and on reverse side upon any matter which in your opinion needs explanation

PERFORMANCE FACTORS	PERFORMANCE GRADE				
	Far Exceeds requirements of this job	**Exceeds** requirements of this job	**Meets** requirements of this job	**Partially Meets** requirements of this job	**Does Not Meet** requirements of this job
QUALITY OF WORK Accuracy Economy of Materials Economy of Time (his own and others) Neatness Thoroughness	Consistently superior ☐	Sometimes superior ☐	Consistently satisfactory ☐	Usually acceptable ☐	Consistently unsatisfactory ☐
QUANTITY OF WORK Productive Output	Consistently exceeds requirements ☐	Frequently exceeds requirements ☐	Meets requirements ☐	Frequently below requirements ☐	Consistently below requirements ☐
DEPENDABILITY Follows Instructions Judgment Punctuality and Attendance Safety Habits	Consistently dependable ☐	Dependable in most respects ☐	Ordinarily dependable ☐	Frequently undependable ☐	Consistently undependable ☐
COMPATIBILITY Attitude Towards the Company Attitude Towards Supervision Co-operation with Fellow-Employees	Inspires others to work with and assist co-workers ☐	Quick to volunteer to work with and assist others ☐	Generally works well with and assists others ☐	Seldom works well with or assists others ☐	Does not work well with or assist others ☐

COMMENTS __________

FIG. 8.2. Example of a multiple-step rating scale. This is a typical rating scale used in industry.

TABLE 8.2

Thirty-five Rating Items Used by Eighteen Companies

A check indicates that the item appears on the company's merit-rating chart. A number following the check indicates the weight assigned to the item. If no numbers appear it is intended to weight all items equally.

	NUMBER OF COMPANY																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
Quality		√8	√	√	√12		√50	√20				√	√	√	√10		√	√50	12
Quantity	√	√6	√	√	√10		√50	√25				√	√	√			√		11
Co-operation	√	√6	√	√			√50	√ 8	√50	√			√	√	√10				11
Initiative	√	√6	√	√	√10	√	√50	√ 6	√50	√	√								11
Dependability	√	√6	√			√		√ 9	√50	√				√	√10		√		10
Personality	√			√		√	√50		√50	√	√		√		√10	√			10
Health	√	√4	√	√	√ 4	√	√50		√50			√				√			10
Safety	√	√6	√		√ 8		√50	√ 7	√50				√	√					9
Industry		√6	√	√	√10	√	√50			√	√				√10				9
Versatility		√6	√	√		√	√50	√15					√				√		8
Leadership	√	√6	√	√	√10			√ 5	√50		√								8
Judgment	√	√4		√			√50		√50	√					√10				7
Intelligence	√	√6	√		√12					√	√					√			7
Attendance		√4	√		√ 4			√ 5					√					√10	6
Knowledge of job			√	√			√50			√	√	√							6
Potentiality	√			√						√	√			√					5
Habits		√4			√ 6	√										√		√10	5
Years of service					√			√15				√						√30	4
Loyalty	√	√4				√													3
Ability to plan	√					√			√50										3
Enthusiasm		√4				√				√									3
Trade skill		√6			√10		√50												3
Technical knowledge		√4	√	√															3
Dependents								√10				√				√			3
Punctuality		√4	√																2
General rating	√								√50										2
Tact	√					√													2
Suggestiveness	√			√															2
Knowledge of costs	√				√ 4														2
Fairness	√																		1
Knowledge of product	√																		1
Knowledge of equipment	√																		1
Knowledge of company policies	√																		1
Appearance						√													1
Place of residence												√							1
Number of rating items	21	19	15	14	13	12	12	11	10	10	7	7	7	6	6	5	4	4	
Method of rating (C = committee; S = supervisor)	S	—	C	S	S	S	C	C	C	—	S	—	S	S	S	—	S	—	
Explanatory phrases used on rating scale	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	No	No	No	

Source: R. B. Starr and R. J. Greenly, "Merit Rating Survey Findings," *Personnel Journal,* 1939, **17,** 378–384.

table. The number of traits included varies from twenty-one in the case of Company 1, which is analyzed in the first column, down to four traits in the case of the companies analyzed in the last two columns. The median number of traits used is ten; therefore, if any virtue lies in considering the practice of the typical or average industry as indicating the correct practice to follow, ten traits would be the proper number to list in a merit-rating chart. Probably, however, little virtue lies in attempting to fit a merit-rating chart to the average of a typical industry; it would, then, be better in any specific instance to adapt the chart to the industry intending to use it than to make it conform to a given set of characteristics, however representative this set of characteristics might be of industry in general.

There is some reason to believe that with the increasing use of merit-rating charts the tendency will be to rate fewer traits rather than more. The primary reason for this belief is to be found in the existence of the "halo" effect, a phenomenon almost always found in psychological ratings. This effect will be discussed in detail on p. 222.

Employee Comparison Systems. Where the rating scales provide for rating against some defined standard (scales on each of several traits), the employee comparison systems provide for rating employees in comparison with each other. With conventional rating scales there frequently is a tendency for the raters to pile up the ratings at one end of the scale, frequently at the higher end. To the extent that this occurs, the results are of limited value since they do not differentiate adequately among the employees. The employee comparison systems avoid this problem completely, since the ratings of employees are relative to each other. There are three principle variations in the methods of comparing employees with each other.

Rank-order system. With this method the rater simply ranks his employees, and each employee's rating is then determined by his position in rank. If a system involving several traits is used, the rankings should be made separately for each trait. The actual process of ranking is usually facilitated by the use of small cards containing the names of men to be ranked. The cards may be arranged and rearranged by the supervisor as he ranks the men on each of the traits considered.

Paired comparison system. Some recent work on the application of a standard psychological method, paired comparisons, to the problem of industrial merit ratings has shown very favorable results.[11] This method ordinarily involves rating the employees on only a single general trait—over-all ability to do their present job. It may, however, be applied separately to more traits or characteristics if that is desirable. Cards or slips of paper are prepared so that each contains the names of two of the men who are to be rated. In this manner, each ratee is paired with every other one. The rater simply checks the name of the man who is considered better on the characteristic on which the men are being compared. The number of pairs of names involved when this system is used in the form described is given by:

$$\text{No. of Pairs} = \frac{N\,(N-1)}{2}$$

where N is the number of men to be rated. If 20 men are involved, the number of pairs is thus $\frac{20(20-1)}{2}$, or 190. But if 100 men are involved, the number of pairs, $\frac{100(100-1)}{2}$, or 4,950, is obviously far too great to permit expedient use of the system. Two solutions have been proposed to the problem that arises when the number of pairs becomes too unwieldy to handle. One method is to divide the original group into a number of smaller groups and set up the pairs for the smaller groups.[12] This procedure not only markedly reduces the number of pairs, but makes more convenient the rating of men in each subgroup by one or more supervisors who are much closer to the men on the job than the supervisor of 100 or more men could possibly be. Scales have been developed so that ratings of men in several subgroups, which may not be equal in size, can be converted to a common base.

A second procedure that may be used when division into smaller groups is not practicable and if there is good reason to believe that the rater is sufficiently familiar with the work of each ratee on the job so that the rating will be reasonably valid, is to extract from the table of all possible pairs of ratees a patterned sample of pairs. It has been shown that the ratings obtained from such a patterned sample correlate approximately .93 with ratings obtained from the complete matrix of pairs.[13]

---

[11] C. H. Lawshe, Jr., N. C. Kephart, and E. J. McCormick, "The Paired Comparison Technique for Rating Performance of Industrial Employees," *Journal of Applied Psychology*, 1949, 33, 69–77.

[12] *Ibid.*

Forced distribution system. This method has been used particularly where a department is quite large—say forty or more employees—and where management does not deem it advisable to set up the cards for the paired comparison system. To meet this situation, the forced distribution system provides that employees be rated on one or more scales, with approximate percentages of employees stipulated for each scale location. Thus, on a scale of "performance on present job," the following percentages might be used:

Lowest 10% *Next* 20% *Middle* 40% *Next* 20% *Highest* 10%

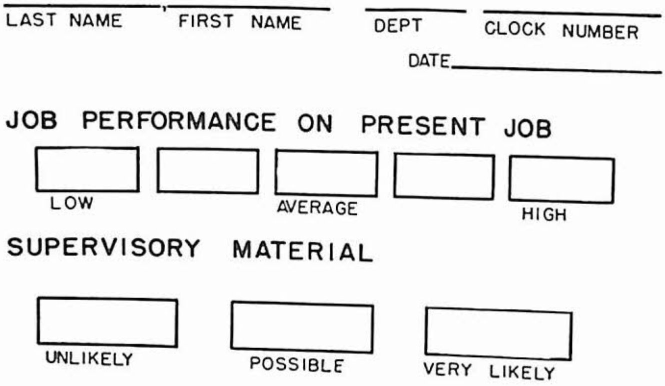
The use of a forced distribution of this type prevents some supervisors from using only the high part of the scale, others from using another part exclusively. The percentages frequently are used as guideposts rather than as rigid rules.

The forced distribution system can be used when rating a number of traits, if that is desired; in this case, the employees may be separately rated on each trait. The evidence from statistical studies such as those reported on pp. 224–226, however, is rather strongly in favor of rating employees on only a few traits. On the basis of this evidence, a simple, two-item, merit-rating card has been developed. On this card, reproduced in Fig. 8.3, each employee is rated on "job performance on present job" and "supervisory possibilities." In the training program preceding the use of these cards the raters are acquainted with the percentage distribution that should be used as a guidepost in rating. It will be noted that no verbal descriptions are printed under the five boxes on the card in Fig. 8.3. Verbal descriptions are not used because, if they were used, the description under the "Low" box would necessarily have to be rather unfavorable, and many supervisors object to placing a man in the low category when it is qualified by an unfavorable comment.

A second rating on "supervisory possibilities" is also called for by the card shown in Fig. 8.3. The use of forced distribution in rating

---

[13] E. J. McCormick and J. A. Bachus, "Paired Comparison Ratings. I. The Effect on Ratings of Reductions in the Number of Pairs," *Journal of Applied Psychology* 1952, **36**, 123–127.



LAST NAME, FIRST NAME

DEPT

CLOCK NUMBER

DATE____________

JOB PERFORMANCE ON PRESENT JOB

☐ ☐ ☐ ☐ ☐

LOW AVERAGE HIGH

SUPERVISORY MATERIAL

☐ UNLIKELY

☐ POSSIBLE

☐ VERY LIKELY

FIG. 8.3. Form used in merit rating employees by a forced distribution system.

this trait is not recommended, because in some departments there may be no employees who are potential supervisors, while in other departments several may have potential supervisory ability.

In using a two-characteristic rating system of this type, the two ratings should not be added together to give a single composite rating. To do so obscures the significance of the individual ratings.

Check List Rating Systems. Check list rating systems embody a third basic method of rating employees. The rater is provided with descriptive statements of employee "behavior," and he is asked to indicate, in one way or another, those statements which are, or are not, descriptive of the employee in question. Thus, the rater tends to be more of a *reporter* of the *work behavior* of employees than an evaluator of their performance or their personal characteristics.

These methods are difficult and time-consuming to develop, so generally they can be justified only where they are to be used on a wide basis. They offer some advantages, however, that warrant their serious consideration for circumstances where they would be widely used.

Weighted check list. This type of merit rating system incorporates a listing of statements of work behavior such as those shown in Table 8.3. This particular check list has been prepared for rating bake shop managers. In rating an employee, the rater simply checks each statement he considers to be descriptive of the work behavior of the individual. It should be added that the scale values shown are *not* on the rating form, but are used later in deriving the employee's rating.

A brief description will be given of the procedures used in developing such a check list. First, a large number of statements are written that describe, in relatively objective terms, various aspects of work behavior, ranging from those that are desirable to those that are undesirable. Next, these statements are "judged" by a number of "experts" on the *degree* to which they are considered to indicate favorable or unfavorable behavior. This is done by the method of "equal-appearing intervals," as described by Thurstone,[14] in which the judges classify the statements in categories (usually 7, 9, or 11) ranging from those they consider to be extremely favorable to those they consider to be extremely unfavorable.

These judgments are then summarized and analyzed, in order to identify the statements that are most reliably judged. Table 8.4 illustrates three hypothetical examples as "judged" by ten judges. Statement A was placed in the same category (4) by all ten judges, and is therefore a highly stable item. Statement B, for which the judgments are scattered a great deal, is a very unstable item. Statement C has moderate stability.

The scale value assigned to an item usually is the mean or median of the categories in which the statement has been classified by the several judges. A possible fallacy in the use of the median for this purpose has been pointed out by Jurgensen.[15] The fallacy depends upon the fact that for some individuals many more items are checked in the top half of the scale than in the lower half. As a substitute he proposes that median values be replaced by positive and negative values obtained by subtracting the mid-value of the scale (4, 5, or 6 depending on whether there are 7, 9, or 11 categories in the scale) from each median. The merit rating score for each individual would then be the algebraic sum of these revised weights for the items checked as applying to the individual.

In a study reported by Uhrbrock,[16] five groups of judges scaled

---

[14] L. L. Thurstone, "Attitudes Can Be Measured," *American Journal of Sociology*, 1928, 33, 529–554.

[15] C. E. Jurgensen, "A Fallacy in the Use of Median Scale Values in Employee Check Lists," *Journal of Applied Psychology*, 1949, 33, 56–58.

[16] R. S. Uhrbrock, "Standardization of 724 Rating Scale Statements," *Personnel Psychology*, 1950, 3, 285–316.

TABLE 8.3

Examples of Items from Weighted Check List Merit Rating for Bake Shop Manager

Item	*Scale Value*
His window display always has customer appeal.	8.5
He gives his employees the reasons for his decisions.	6.7
Products dropped on the floor are sometimes sold.	1.4
He always gets his reports in on time.	7.9
He rarely figures the costs of his products.	1.0
He belongs to a local merchants' association.	4.9
His bakers tend to pass some of their work off onto him.	3.1
He seldom forgets what he has once been told.	7.6
He occasionally runs a selling contest among his salesgirls.	6.8
He is slow to discipline his employees even when he should.	1.9
He does not anticipate probable emergencies.	2.4
His weekly and monthly reports are sometimes inaccurate.	4.2
He should take more interest in merchandising.	3.5
He is slow at making decisions.	4.4
His bakers do not respect him.	0.8
His sales per customer are relatively high.	7.4
He has originated one or more workable new formulas.	6.4
No baking is done in his shop after 12 noon.	0.6
He encourages his employees to show initiative.	8.1
He knows how but can't teach others.	2.5
His shop is unusually neat and clean.	8.3
He often has vermin and insects in his shop.	0.8
His salesgirls sometimes fail to use wax paper to handle goods in the sales room.	3.2
He pays little attention to his customers.	1.6

Source: Adapted from E. B. Knauft, "Construction and Use of Weighted Check-List Rating Scales for Two Industrial Situations," *Journal of Applied Psychology*, 1948, **32**, 63–70.

TABLE 8.4

Hypothetical Examples of Judgments by Ten Judges on Three Check List Rating Statements*

Statement	*Rating Category*								
	Unfavorable							*Favorable*	
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
A				10					
B		1	2	1	3	2	1		
C	2	5	3						

*The number entered under each category is the number of judges who placed the statement in that category.

a large number of general rating scale statements. The groups included 20 Foremen, 10 Technical Assistants, 24 University Students, and groups of 10 and of 25 University Faculty members. Correlations of about .96 and .97 among these groups indicate that the rather small groups of judges can give reliable ratings, and that judges of different backgrounds agree quite consistently regarding the "values" of rating scale statements. A few examples from the ratings by one of these groups—the foremen—are shown in Table 8.5. The "variability" column gives an index of the relative variability of the judgments (the smaller this value, the greater the degree of consistency in the judgments).

TABLE 8.5

Scale Values of Selected Check List Rating Scale Statements as Derived from Judgments of Twenty Foremen

Statement	*Scale * Value*	*Variability †*
6. Makes the same mistakes over and over.	14	33
34. Is a clock-watcher.	22	179
67. Generally looks for the easy work.	27	231
83. Can't seem to get the hang of things.	29	79
152. Is somewhat ill-natured.	36	54
171. Conduct borders on insubordination.	36	535
238. Can do good work if he (she) tries.	49	79
251. Is more interested in ideas than people.	57	541
257. Never quits ahead of time.	64	124
277. Is orderly in work habits.	69	219
297. Makes friends with others easily.	72	153
307. Can stand criticism without feeling hurt.	74	194
377. Gets production out in less than average time.	79	109
447. Submits clear, understandable reports.	84	283
498. Can concentrate under difficult conditions.	90	190
519. Is at ease in any situation.	98	179
539. Merits the very highest recommendation.	107	51

* Scale values have been multiplied by 10 to avoid decimal values.

† The "variability" is the square of the standard deviation of the placements multiplied by 100. It can be interpreted as a *relative* index of the consistency of judgments.

Source: Uhrbrock, *op. cit.*

Forced-choice check list. During World War II the Personnel Research Section of the Adjutant General's Office developed this unique method to rate Army officers.[17] In the form in which it was originally developed, this rating system consists of a number of groups of four words or phrases that describe a good (or poor) officer. Each group of four statements includes two "favorable" statements and two "unfavorable" statements. The rater checks the statement that is *most* like the person being rated, and the one that is *least* like him. In later developments of the method various modifications have been used, with variations in the number of statements in a block and in the number and type of rater response required. Some variations studied by Berkshire and Highland[18] included blocks with two, three, four, and five statements, with certain groups including only favorable or unfavorable statements, but not both. An example of a block of forced-choice statements is given below. This particular illustration is from a rating form used for rating Air Force instructors; all of these statements are "favorable."[19]

---

[17] E. D. Sisson, "Forced Choice—The New Army Rating," *Personnel Psychology*, 1948, 1, 365–381.

a.) Patient with slow learners.
b.) Lectures with confidence.
c.) Keeps interest and attention of class.
d.) Acquaints classes with objective for each class in advance.

The selection of the statements for each block is based on extensive preliminary research to determine the *degree* to which each statement is considered by raters generally to be a "favorable" or "unfavorable" statement (favorability index), and the extent to which the statement, when used in a rating situation, tends to *discriminate* between above- and below-average individuals (discrimination index). The statements are then grouped together on the basis of relatively comparable favorability indexes; the statements above, for example, have been placed in the same block because they have approximately *equal favorability* indexes. The items placed in a block, however, *differ* in their *discrimination* indexes, usually only one in a block being a discriminative item.

A major advantage of this method is that a rater, attempting (consciously or unconsciously) to rate a man higher, or lower, than the man's true worth, has no way of knowing which of the statements to check to raise (or lower) the man's rating from what it should be. There are, however, certain disadvantages to the method. In the first place, since the rater does not (and should not) know how the final rating values are derived, he may resent the system as a whole, and may therefore not give it his wholehearted support. In addition, the method does not lend itself readily to counseling with the employee.

---

[18] J. R. Berkshire and R. W. Highland, "Forced-choice Performance Rating—A Methodological Study," *Personnel Psychology*, 1953, 6, 355–378.

[19] Berkshire and Highland, *op. cit.*

Of the various versions of this method, Berkshire and Highland[20] found that the form with four favorable statements from which the rater was to select the two most like the ratee, had advantages over other forms. It was the most resistant to bias, it yielded consistently high validities under various conditions, it had adequate reliability, and was one of the two forms best-liked by the raters.

In some situations, the forced-choice method is used in combination with other rating procedures. This has the advantage of making the combined rating procedures useful for more purposes.

Critical Incident Technique. This method of merit rating, developed by Flanagan and Burns[21] provides for the recording, by supervisors, of critical "behaviors" on the part of employees. Whenever an employee does something that is especially noteworthy, or especially undesirable ("critical" to either good or poor performance), a notation is made in the employee's record. These "critical behaviors" usually are classified into certain categories, such as those given in Table 8.6. When this rating system is used, supervisors note and record all "critical" instances of on-the-job behavior falling in any of the categories listed in Table 8.6.

Although this method does not readily lend itself to objective quantification, it offers a strong advantage for purposes of employee counseling because it provides the supervisor with a record of "specifics" to discuss with the employee.

SOME DANGERS OF MERIT RATING

In the previous discussion of merit ratings there have been implications of some of their limitations or dangers. Most of these limitations are attributed to certain rating tendencies on the part of raters. These tendencies are most pronounced with conventional rating scales. In fact, some of the other rating methods have been developed primarily for the purpose of avoiding or reducing the influence of undesirable tendencies on the part of raters. The discussion of these tendencies will deal largely with conventional rating scales.

---

[20] Berkshire and Highland, *op. cit.*

[21] J. C. Flanagan and R. K. Burns, "The Employee Performance Record: A New Appraisal and Development Tool," *Harvard Business Review*, 1955, 33 (No. 5), 95–102.

TABLE 8.6

Examples of Categories of Types of Critical Incidents

1. Physical condition.
2. Coordination.
3. Checking and inspecting.
4. Arithmetic computations.
5. Learning and remembering procedures and instructions.
6. Judgment and comprehension.
7. Understanding and repairing mechanical devices.
8. Improving equipment and showing inventiveness.
9. Productivity.
10. Dependability.
11. Accepting supervision and organizational procedures.
12. Accuracy of reporting.
13. Response to departmental needs.
14. Getting along with others.
15. Initiative.
16. Responsibility.

Source: J. C. Flanagan and R. B. Miller, "The Performance Record: Handbook for Foremen and Supervisors of Hourly Employees," Science Research Associates, Chicago, Ill., 1955.

The Halo Effect. More than thirty years ago Thorndike[22] pointed out on the basis of experimental evidence that a rater has a constant tendency to rate an individual either high or low in many traits because the rater knows (or thinks) the individual to be high or low in some specific or particular trait. Thorndike called this tendency the "halo" effect. Applied to the industrial situation, Thorndike's statement means that if the supervisor regards an employee as very satisfactory in terms of a trait such as general personality or tact, he is likely to rate the employee high also in such traits as productivity, ingenuity, inventiveness, adaptability, and perhaps many other traits.[23] In other words, it is difficult for any rater—particularly an untrained rater—to isolate and rate separately the various traits that an employee may possess. The use of a merit-rating chart is likely to increase the ability of a foreman to make an analytical judgment; but we know that even under the most favorable conditions the halo effect will be present to some extent and that its results will be most prominent where ratings have been made by those unfamiliar with its very existence.

---

[22] E. L. Thorndike, "A Constant Error in Psychological Ratings," *Journal of Applied Psychology*, 1920, **4**, 25–29.

[23] A discussion of the halo effect in the industrial application of merit rating is given by R. S. Driver, "A Case History in Merit Rating," *Personnel*, 1940, **16**, 137–162.

The halo effect can be minimized in several ways. If a rating scale system is used, it is usually advisable to follow the suggestion of Stevens and Wonderlic[24] and have each supervisor rate all his men on one trait before going on to the second trait, on the second trait before going on to the third, and so on. Since this method causes the supervisor to think of all his men in connection with a given trait alone rather than to think of each man as a whole, the effect of this general change in point of view is a reduction of the halo effect. This effect can also be minimized by arranging the chart itself so that the desirable end of some traits is on the right-hand side of the scale whereas the desirable end of other traits is on the left-hand side of the scale. This procedure prevents a supervisor from checking down a column on the right-hand side for a generally desirable employee, or down the left-hand side for a generally undesirable employee.

The operation of the halo effect in an actual set of ratings is shown in Fig. 8.4. This figure reveals graphically, for eighteen men selected at random from a large industrial organization, the relationship between ratings on over-all job performance, accuracy, safety, and co-operation, on the one hand, and on the other, an over-all merit rating that included the four traits mentioned above along with eight others that need not be identified. The eighteen men are arranged in rank order from left to right according to their over-all merit rating. These ratings are indicated by the heavy black line. The ratings of the eighteen men on the four traits are indicated by the remaining lines. It will be seen from an inspection of Fig. 8.4 that those individuals who were rated high in any given trait, say co-operation, were also rated about equally high on all of the other traits shown. Of course, it may theoretically be possible that the individual high in one trait is actually high in all of the other traits, but it seems much more likely that the halo effect, rather than any real relationship among the traits, is operating here.

Further evidence of a more extensive statistical sort of the existence of the halo effect has been obtained from a factor analysis of approximately 1,100 ratings selected at random from a plant in which

---

[24] S. N. Stevens and E. F. Wonderlic, "An Effective Revision of the Rating Technique," *Personnel Journal*, 1934, 13, 125–134.

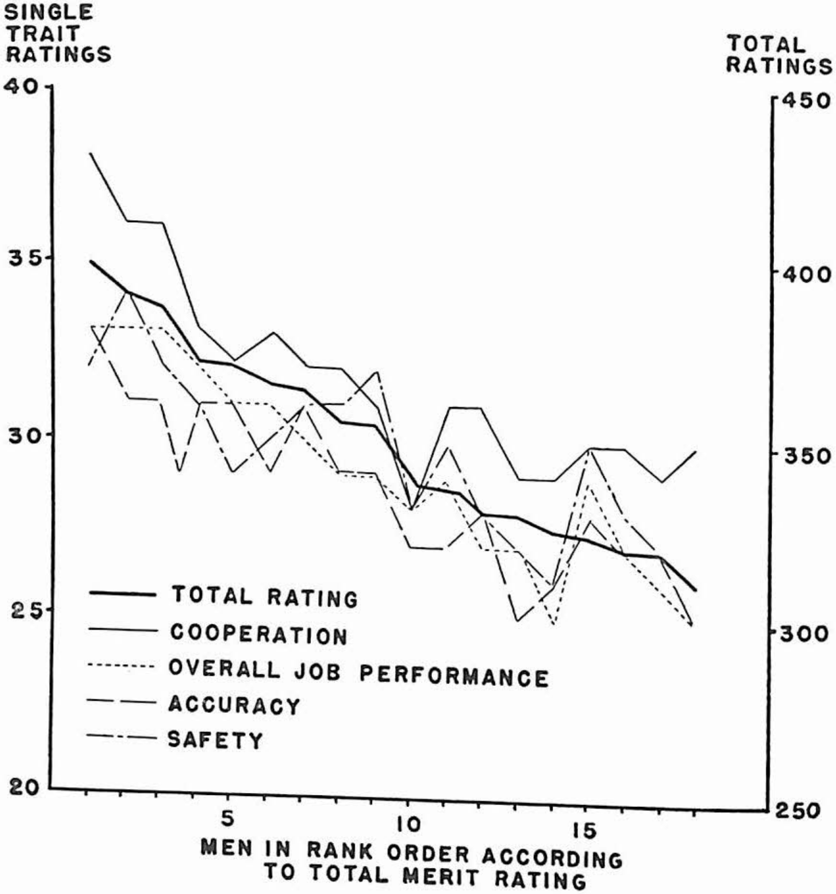


FIG. 8.4. Operation of the "halo effect" in industrial merit rating.

about 9,000 employees had been rated.[25] The chart used involved ratings on twelve separate traits. These traits and the intercorrelations between the ratings for each pair of traits are shown in Table 8.7. Such a correlational matrix containing correlations in general as high as those in Table 8.7 is sufficient in itself to indicate that a common factor is operating to influence most, if not all, of the ratings. The results of a factor analysis of these ratings quite definitely justify this judgment. Factor analysis is a statistical technique that reduces

---

[25] Edwin Ewart, S. E. Seashore, and Joseph Tiffin, "A Factor Analysis of an Industrial Merit Rating Scale," *Journal of Applied Psychology*, 1941, 25, 481–486.

TABLE 8.7

Intercorrelations of the Traits in a Twelve-Item Merit-Rating Scale

Traits	1. Safety	2. Knowledge of Job	3. Versatility	4. Accuracy	5. Productivity	6. Over-all Job Performance	7. Industriousness	8. Initiative	9. Judgment	10. Co-operation	11. Personality	12. Health
1. Safety												
2. Knowledge of Job	.61											
3. Versatility	.52	.81										
4. Accuracy	.63	.85	.80									
5. Productivity	.55	.79	.72	.81								
6. Over-all Job Performance	.60	.82	.80	.67	.86							
7. Industriousness	.49	.78	.71	.80	.86	.85						
8. Initiative	.54	.78	.78	.78	.80	.83	.82					
9. Judgment	.62	.80	.82	.84	.81	.88	.84	.86				
10. Co-operation	.61	.67	.68	.74	.81	.80	.80	.72	.76			
11. Personality	.55	.67	.63	.70	.73	.74	.67	.72	.75	.80		
12. Health	.25	.52	.50	.84	.45	.60	.53	.77	.43	.52	.71	

a set of measurements (such as test results or merit ratings) to the minimum number of basic variables or factors that will account for the variations in the original data. The factor analysis of the merit ratings revealed three basic factors, factor loadings for which are given in Table 8.8. In interpreting the importance of the different merit-rating items as they enter into each of the three factors, we should bear in mind that the amount of the factor loading indicates the extent to which that item is related to the factor in question. An inspection of the merit-rating items entering into Factor I reveals that over-all job performance is loaded the most heavily; productivity, second; industriousness, third; judgment, fourth; and so on down the list. If we were to attempt to name Factor I, probably the best name available would be "ability to do the present job." In other words, this factor seems to be a job-performance factor. Since the factor loadings on all of the traits entering into Factor I are high,

TABLE 8.8

Factor Loading After Rotation

Trait	Factor Loadings[†]			h^2[*]
	Factor I	Factor II	Factor III	
1. Safety	.633	.245	—.039	.462
2. Knowledge of job	.841	.337	—.002	.821
3. Versatility	.795	.325	.030	.739
4. Accuracy	.826	.448	.297	.971
5. Productivity	.913	.039	—.042	.837
6. Over-all job performance	.961	—.080	—.064	.934
7. Industriousness	.906	.009	.007	.821
8. Initiative	.887	.094	.252	.859
9. Judgment	.897	.265	—.036	.876
10. Co-operation	.881	—.025	.042	.779
11. Personality	.815	.009	.259	.731
12. Health	.573	—.016	.836	1.027

[*] Communalities.

[†] A factor loading is the coefficient of correlation between the specific variable and the factor involved.

compared with the loadings on Factors II and III (except "health" in Factor III), we infer that this factor is by far the most important of the three in determining the over-all merit rating and that, since all of the heavily loaded traits entering into Factor I deal rather specifically with ability on the job, it would seem safe to identify this factor as a job-performance factor.

Factor II contains no elements so heavily loaded as are those entering into Factor I. Those most heavily loaded in Factor II are accuracy, knowledge of the job, and versatility. Judgment and safety are the fourth and fifth, respectively, but they are not so highly related to Factor II as are the first three traits named. If we attempt to name Factor II, a name must be found that combines the five traits as they are all related to the unnamed factor. Since none of the loadings for this factor is high in comparison with those for Factor I, we must conclude that none of the twelve items of the rating scale "taps" this factor very well. We are faced then with the task of naming this factor although we know only the following facts: (1) accuracy correlates with it .44, (2) knowledge of the job correlates with it .34, (3) versatility correlates with it .33, (4) judgment correlates with it .27, (5) safety correlates with it .25, and (6) no one of these items from the rating chart is a satisfactory name for the factor because no one of them correlates highly with the factor.

From this evidence it would seem that Factor II may deal with the *quality* of performance on the job. Skill and accuracy are also suggested as possible names. Since knowledge of the job has the second highest factor loading, perhaps promotability or likelihood of being promoted would be an appropriate identification of this factor. Whatever we may choose to call it, certainly this second factor is not the same as Factor I, and it does not have as much influence on the total ratings, under the system that had been in use when these merit ratings were obtained, as does Factor I.

Factor III was found to be significantly loaded with only one element, namely, health. In other words, this factor was related to the health ratings and to little else. A further analysis of this factor supports the inference that the "health" ratings were grossly unreliable. Thus, the available evidence regarding this factor provides sufficient grounds for disregarding it. A further reason (if one were needed) for disregarding the health ratings would be that an employee's health is, of all things, something that should be determined scientifically by a physician and not by the judgment of his foreman or supervisor.

The logical conclusion one would reach from the foregoing discussion is that the merit-rating chart in question should be reduced to only two traits and that these traits should be identified to cover as well as possible the two factors that the analysis revealed. This conclusion is the basic reason for the use of some of the employee comparison systems, which usually provide for ratings on some broad aspect of performance such as "over-all job performance." Such a reduction in number of traits rated involves no sacrifice in either reliability or accuracy. It may well be, however, that certain other advantages that accrue from merit rating—such as the increase in the analytical viewpoint of the foreman and the identification of the employee who needs help or training—can be better achieved by a more detailed type of merit-rating chart. The advantages of the analytical approach, however, sometimes can be obtained by the use of two or more methods in combination. The decision on which system is to be used must be made by each company in the light of the evidence and after a careful consideration of the needs of the particular plant.

A similar factor analysis study by Kelly[26] of a fourteen-item graphic scale for rating pilot competency showed that three factors (skill, judgment, and emotional control) accounted for all intercorrelations among the fourteen original items.

It should be stated that the presence of only two or three factors in the studies cited may be a reflection of inadequate training of the raters in the general principles of rating. A study by Driver[27] of a similar set of ratings obtained by raters who had been given seven hours of intensive training in rating methods revealed much lower intercorrelations than those shown in Table 8.7.

The general conclusion with regard to the halo effect is that unless supervisors have been trained very carefully they may allow their general impression of an employee, or their evaluation of one of his traits, to influence their ratings on other traits as well.

Constant Error. Some raters tend to be lenient in making ratings, thereby giving many employees "high" ratings. Other raters tend to be "tough," and give primarily "low" ratings. Unless a suitable correction is made for such tendencies of different raters, ratings of different men that have been made by different raters cannot reasonably be compared. For example, in using a twelve-factor merit rating system, each factor having a maximum value of 50 points, the ratings made by one supervisor averaged 405 points, while the ratings made by another supervisor (on a presumably comparable group of men) averaged 295 points. The ratings by these two supervisors are not comparable unless they are converted to a common base.

Influence of Other Factors on Ratings. It sometimes happens that raters are influenced, unknowingly, by various factors, in some cases by factors that actually should not have such influence. Among such factors are department, job, age, and length of experience.

Departmental differences. Very frequently the merit ratings turned in from different departments in a given plant differ markedly from one department to another. This discrepancy may be due in part to actual differences in the merit of employees in the various departments, but it may also be due in part to differences in standards or interpretation of the merit-rating scale among the departments.

---

[26] E. L. Kelly, "The Development of a Scale for Rating Pilot Competency," *CAA Div., Res. Rep. No. 18, 1943, Publ. Bd. No. 50297* (Washington, D.C., United States Department of Commerce, 1947).

[27] Unpublished study by R. S. Driver, Atlantic Refining Co., Philadelphia.

Whatever may be the cause, when such differences do occur it is usually desirable to evaluate a given employee's rating in terms of the other ratings from the department in which he is working rather than in terms of the ratings obtained from the plant as a whole. This situation is graphically illustrated in Fig. 8.5, which shows the

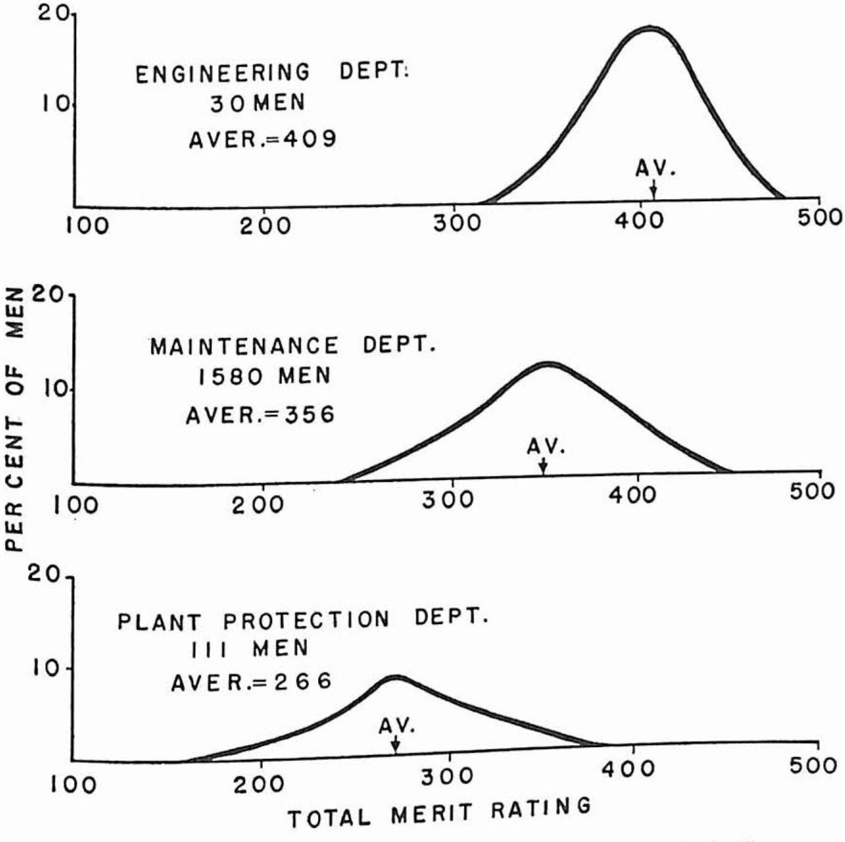


FIG. 8.5. Differences in merit ratings among departments in a steel mill.

distributions of total merit ratings obtained from three departments of a plant made up of fourteen departments and employing approximately 10,000 men. The three departments selected for illustration are engineering, maintenance, and plant protection. These three have been selected to show the marked differences that may be found in ratings from one department to another. The difficulty of interpreting the significance of a merit rating without reference to the department from which it was obtained may be readily seen from inspecting Fig. 8.5. For example, a merit-rating score of 350 would be a very low rating for a man in the engineering department. The same rating would be approximately average for a man in the maintenance department, and it would be very high for a man in the plant-protection department. Since one major function of a merit rating is to indicate how well—in relation to other employees—an employee is doing his present job, a fair and reasonable basis for comparing ratings of different men must be employed. When ratings differ markedly from one department to another, evaluation of any rating should be in terms of the department from which it was obtained.

Job differences. Another source of possible difficulty closely related to the matter of departmental differences is the variation in rating often found from one job to another. When employees on any given job are consistently given higher merit ratings than are employees on other jobs, such job differences should be considered in evaluating the rating of any given employee. Fig. 8.6 illustrates this situation for 51 jobs in a sheet and tin mill. The 51 jobs studied are arranged in order from the one receiving the highest average rating (tinner) at the top to the one receiving the lowest average rating (opener and examiner) at the bottom. It will be noted that there is a variation from 275 to 385, or 110 points, in average merit rating from the lowest to the highest average rating. From these differences it is clear that a rating of 300 is very high for an employee who is on one of the jobs located near the bottom of the list, but that 300 is very low for an employee on one of the jobs near the top of the list. The implication of this fact is that the merit rating of an employee should be evaluated in relation to the ratings of other employees on the same job or on jobs where the employees are given approximately the same average merit rating.

Age of employees. The age of an employee is another factor that is often related to the rating he receives. In one set of approximately 9,000 merit ratings obtained from a single industrial plant, the relationship between total rating and age is graphed (Fig. 8.7). This figure shows that an employee between the ages of thirty and thirty-five, other things being equal, is likely to get a higher merit rating

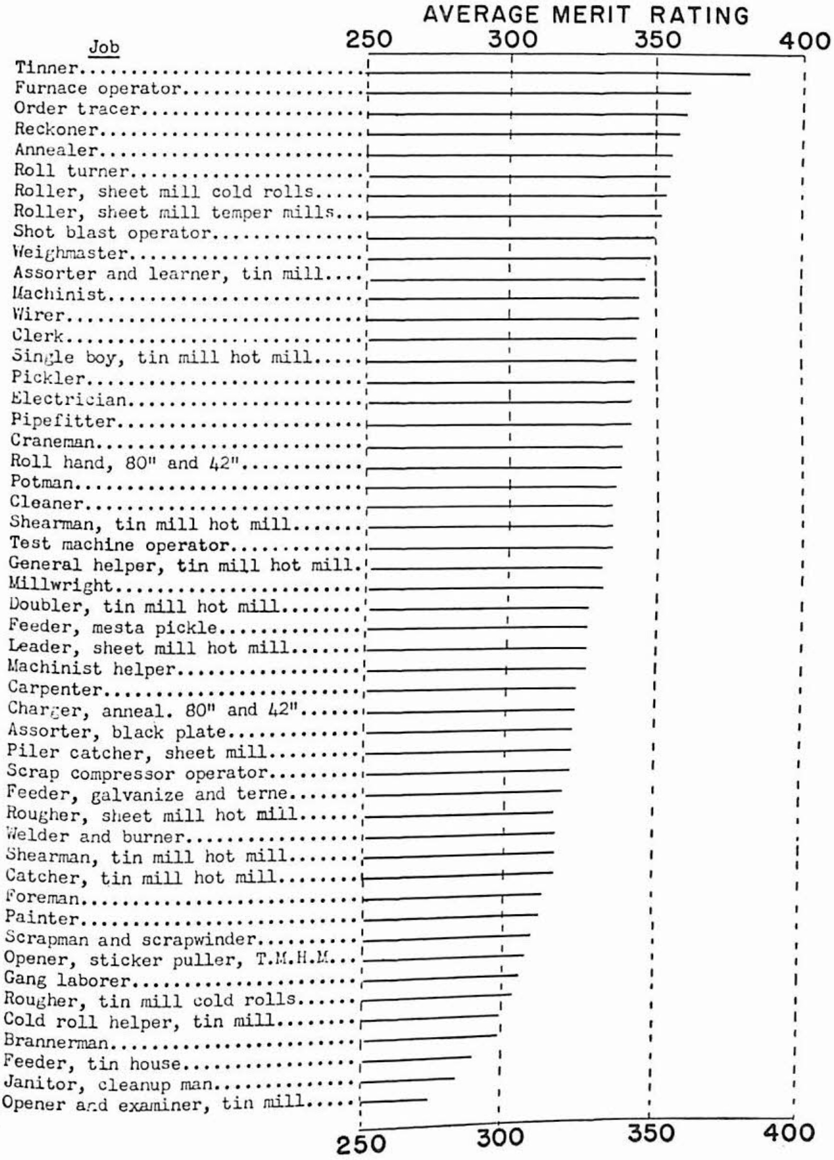


FIG. 8.6. Differences in average merit rating of employees on 51 jobs in a sheet and tin mill.

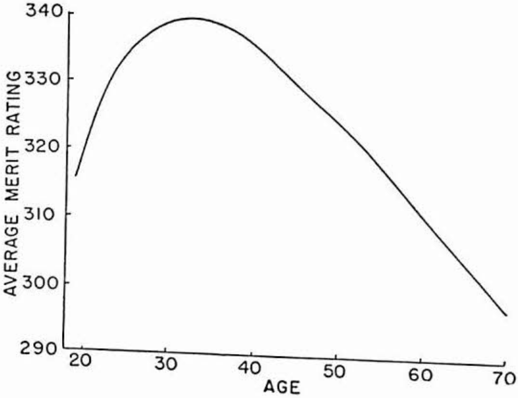


FIG. 8.7. **Relation between age and average merit rating for 9,000 steel workers.**

than an employee who is either older or younger. Fig. 8.7 also shows that after this high point is reached a progressive decrease in merit rating takes place, as age increases. It is clear from this chart that an employee with a merit rating of 330 is definitely below average if he is in the age range from thirty to thirty-five, but that he is well above average if he is in the age range from fifty-five to sixty.

Other factors. The preceding factors that should be considered in evaluating the merit rating of employees have been mentioned as illustrative of the kind of factors that have been found in a number of investigations to be related to merit ratings. Still other factors exist that may affect merit ratings in any given plant. Fig. 8.8, for example, illustrates the relationship found in one plant between merit ratings and length of service on the present job. In definite contradiction to the opinion of management before these results were obtained, this chart shows a progressive lowering of merit ratings as the length of service on the job continues. In other words, the longer a given employee remains on a certain job, the lower the merit rating he is likely to receive from his supervisor. The probable explanation for this relationship is that only those employees who are unqualified for promotion or transfer to another or more important job are likely to remain on their present jobs for a long period of time. Or perhaps, in this particular plant, the employees with longer service on their present jobs are in general older employees and may be unable to do the work as effectively as younger

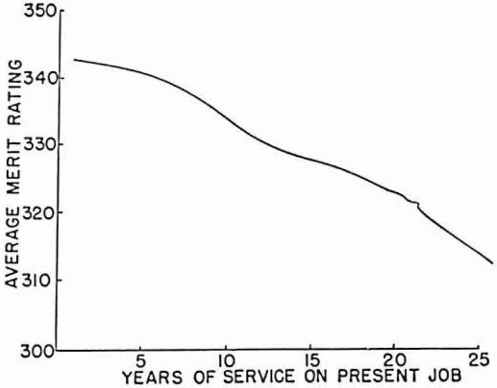


FIG. 8.8. Relation between years of service on present job and average merit rating for 9,000 steel workers.

men. Whatever may be the cause of the relationship, it is clear that it exists and that it should, therefore, be considered in evaluating a given employee's rating. A rating, such as 335, that might be well below average for an employee of short service on the job, would be well above average for an employee who has been on the job fifteen or twenty years.

Fig. 8.9 graphs the relationship between merit rating and total service in the plant. It is clear that the fluctuation in average merit rating with total plant service is much less marked in the case of this relationship than in the case of the relationship with length of service on the present job. There was no marked or systematic change in total merit rating with total plant service.

It is not meant to imply from the results summarized in the preceding sections that the relationships found in the studies reported exist in exactly this same form in all industries. Indeed, there is every reason to expect that these relationships are not universal and that opposite trends may occur in many plants. Rothe has emphasized that there is no single fixed relationship between merit ratings and length of service.[28] The point to keep in mind is not the universality of the particular trends discussed but the fact that no company can properly interpret the results of the merit ratings in its own plants

---

[28] H. F. Rothe, "The Relation of Merit Ratings to Length of Service," *Personnel Psychology*, 1949, 2, 237–242.

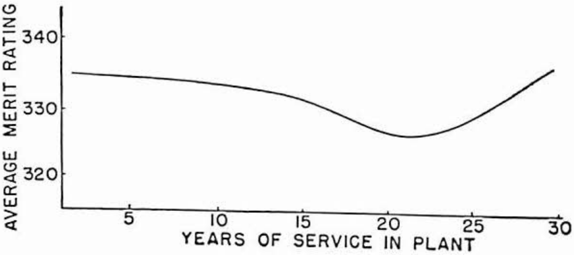


FIG. 8.9. Relation between years of plant service and average merit rating for 9,000 steel workers.

without definite knowledge of what trends and extraneous factors influence these ratings in that particular organization.

Giving Rating Information to Employees. In one set of ratings studied, the numerical values, obtained by adding the points contributed by the twelve items comprising the scale, varied from 150 to 500. The reliability of this set of ratings as discussed in a following section is such that two employees must differ by at least 66 points before it can safely be assumed that a real difference between the two employees exists. In the light of this situation it is unwise to give the employees exact numerical statements of their ratings. One learns early in life that the figure 400 is larger than the figure 399. Therefore, if two employees should receive ratings represented by these numbers, one is likely to feel elated and the other depressed, though there is no reason at all for assuming that such a slight difference represents any real difference between the two employees. It is much safer to divide the range of ratings into four or five categories and to tell each employee only in which category he is located rather than exactly where he stands in that category.

A second problem in giving out numerical ratings arises when an employee compares his rating with a rating given him at some previous time. Suppose, for example, an employee's rating in 1958 is 350. He is not satisfied and makes every effort during the following year to improve his performance on the job. Possibly he attends night school, reads material related to his job, and in other ways makes a sincere effort to improve his rating. At the end of the year the ratings are repeated and he finds that this time his numerical rating is only 345. One familiar with the fact that ratings are none too reliable, even at best, would certainly not consider such a slight decrease from one year to another as indicating a drop in the employee's merit. But the employee himself, if these numerical figures have been given out, is very likely to adopt the attitude, "What's the use?" If he has done everything possible during the year to improve himself, and finds at the end of that time that his rating is lower than it was before, he is unlikely to make any serious attempt in the future to improve his ability. This difficulty can be largely eliminated by giving out only general classifications such as A, B, C, and D. A "B-grade" employee is less likely to experience a change of grade as a result of a few numerical points difference in his rating than if he is classified exactly according to his numerical standing.

ADJUSTING RATINGS FOR UNDESIRED INFLUENCES

We can see from the above discussion that ratings may be influenced by the rater's "constant" error, or may be related to such factors as the job, the department, and age. In order to determine whether there *are* differences in the ratings of employees by rater, job, and department, it is necessary to make an analysis of the ratings such as those previously shown in Fig. 8.5, 8.6, 8.7, 8.8, and 8.9.

Where it is found that ratings *do* differ significantly for the employees in different categories (such as employees rated by different raters, those in different departments, those on different jobs, and those of different ages), there are two possible explanations for the differences. In the first place it is possible that the differences are real differences—that the employees in one category in general *are* better than those in another category. The other possible explanation is that the ratings have been influenced by spurious factors such as raters' tendencies, and that the differences are simply functions of these influences; in this case the systematic differences—say, in the ratings given by two raters—are not true differences. Where there are significant differences in ratings of employees in various rating groups (such as employees rated by two or more raters), the fact of these differences is immaterial *if* the ratings are to be used *entirely within the context of the group*. If the ratings are to be *compared between or among* rating groups, however, it is pertinent to make an informal judgment as to whether there is any justification for believing that the differences are *real* or *spurious*. In the absence of evidence to the contrary, it is probably most justifiable to assume that they are attributable to raters' tendencies or similar factors that should be ruled out.

In such a case the ratings should not be compared across rating groups *unless* they are adjusted for the differences. There are at least two ways of doing this.

Adjusting for Differences in Means. Let us continue to use, as an example, the situation in which the ratings given by different raters differ significantly. In such a case, one solution is to determine the average of the ratings given by each rater, and compute the difference between *each rater's* average, and the average of *all* raters. This difference can then be added to or subtracted from the ratings given by a particular rater, in order to bring his ratings into alignment with those of other raters. This simple adjustment is satisfactory if the *variabilities* of the ratings given by the different raters are about the same.[29]

Adjusting for Differences by Standard Scores. A more systematic method of adjusting for such differences consists of converting all ratings to a common set of numerical values. Some type of *standard score* (comparable score) may be used for this purpose. There are various types of standard scores, such as *z*-scores. Standard scores, including *z*-scores, indicate *relatively* the position of individual cases in a distribution. Such scores are based on deviations of individual cases from the mean *in standard deviation units.* A "standard deviation" is a statistical index of the degree of variability of the cases within a distribution.[30] It is expressed in terms of the numerical values of the original distribution. In a relatively normal distribution, two-thirds of the cases fall within one standard deviation above and below the mean, about 95 per cent are within two standard deviations above and below the mean, and about 99 per cent fall within three standard deviations. Thus, regardless of what the mean of a distribution is, or what the magnitude of its standard deviation, it is possible to express the deviation of any given numerical value in terms of the number of standard deviation units it is above or below the mean. A *z*-score is simply the deviation of a given raw score from the mean expressed in terms of standard deviation units.

---

[29] The variabilities in the ratings can be compared by comparing their *standard deviations*. See Appendix A, pp. 525–527.

[30] *Ibid.*

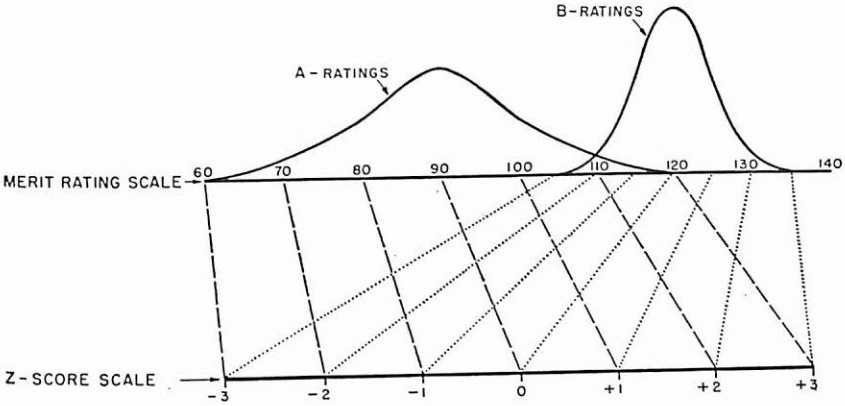


FIG. 8.10. Illustration of conversion of two sets of ratings (A and B) to a common scale of z-scores.

Let us now see how this helps us in comparing the merit ratings of employees as rated by a "tough" rater and as rated by a "lenient" rater. Let us suppose that the distributions A and B of Fig. 8.10 represent, respectively, the total ratings given to their respective groups by rater A and rater B. We can see clearly that a rating of, let us say, 110 by itself, is meaningless unless we relate it to the distribution of which it is a part (it means a very high rating by rater A, and a very low rating by rater B). These two distributions, however, can both be converted to z-scores, as illustrated at certain points on the distributions by the broken lines and dotted lines from the merit rating scale to the z-score scale below it. We can now see that a rating of 110 by rater A would mean a z-score of *plus* 2, but by rater B would mean a z-score of *minus* 2, and that a rating of 100 from rater A would correspond to a rating of 125 from rater B, since both convert to z-scores of plus 1.

Similar conversions can be made for ratings of people from different rating groups, such as for different jobs or departments.

THE "WEIGHTING" OF TRAITS

In the conventional rating scales for merit rating purposes it is customary to rate employees on each of several traits or characteristics. Usually it is considered desirable to weight each of these traits according to what seems to be the relative importance of each for success of employees in that particular organization. Thus one company might weight initiative 20 points and judgment 5 points, whereas another organization might reverse this weighting of items. Table 8.2 shows that of the eighteen charts analyzed, four, or 22 per cent, incorporated some sort of differential weighting for the various items. The remaining charts use either equal weightings or no weightings at all. Either of these methods is presumably intended to give the same importance to the several traits included on the chart.

At least two difficulties arise in connection with weighting items on a merit-rating chart. The first of these is that items are not necessarily weighted equally when all are given the same maximum numerical value, nor are they necessarily weighted in the way intended when a predetermined set of maximum values for each is used. In combining scores—whether they are merit-rating items, test scores, production records, or any other set of values—the scores weight themselves automatically in proportion to their respective variability. Expressed in statistical terms, the scores weight themselves in proportion to their respective standard deviations. Therefore, if the variability of all employees on one rating item, say judgment, is twice as large as the corresponding variability of all employees on some other item, say initiative, a direct combination of ratings for any employee on these two traits is actually weighting the judgment ratings twice as heavily as the initiative ratings. It might, of course, be the wish of management to weight those two items in this proportion, but it is unlikely that the chance weights that creep into a set of ratings as a function of their respective variabilities will weight the various traits in the manner desired by management.

The statistical reason for weights being determined by variability of the ratings is discussed in some detail in Appendix A, p. 527. An example here may further clarify the principle.

Suppose that 1,000 men have been rated on two traits, namely, judgment and initiative. Each man has been rated on each trait on a 50-point scale. Suppose, for the present illustration, that all of the men have received ratings on initiative of between 30 and 35 points. Suppose, further, that the judgment ratings vary from 25 to 45. If we now combine for each man his rating on initiative and his rating on judgment we will obtain a combination rating in which it has often been assumed (because each trait was originally rated on a 50-

point scale) that the two ratings are weighted equally. Under these circumstances, however, the traits are not weighted equally at all. The judgment ratings, which vary over a range of 20 points—from 25 to 45—will have approximately four times as much effect on the total rating as the ratings on initiative, which vary over a range of only 5 points—from 30 to 35. A method of combining the ratings so that they may be weighted equally is described in Appendix A. If this method seems rather complicated, it can only be said that the rating of human personality traits is complicated and is difficult enough to accomplish satisfactorily even when all statistical safeguards are followed.

The fact that unknown weights for the various items on a merit-rating chart not only may but *do* creep in, if not guarded against statistically, is proved by an analysis of the merit ratings of 1,000 employees in a steel mill. The variability of the ratings on each of twelve traits was computed and the standard deviations are shown in Table 8.9.

TABLE 8.9

Standard Deviations of Ratings of Employees in Twelve Traits

Trait	*Standard Deviation*	*Relative Weight*
1. Safety	2.24	1.00
2. Knowledge of job	2.77	1.24
3. Versatility	2.88	1.29
4. Accuracy	2.69	1.20
5. Productivity	2.58	1.15
6. Over-all job performance	2.63	1.18
7. Industriousness	2.96	1.32
8. Initiative	3.08	1.38
9. Judgment	2.68	1.20
10. Co-operation	2.72	1.22
11. Personality	2.51	1.12
12. Health	3.14	1.40

Source: J. Tiffin and W. Musser, "Weighting Merit Rating Items," *Journal of Applied Psychology*, 1942, **26**, 575–583.

The employees had been rated on each of these traits on a 50-point scale, and it had been assumed by the management that this procedure resulted in total ratings that were influenced in an equal amount by each trait. Actually, the health ratings, which had the largest standard deviation, were exercising 40 per cent more effect on the total ratings than the safety ratings, which had the smallest standard deviation. Next in importance were the ratings on initiative, which were exercising 38 per cent more effect than the safety ratings. The relative weights actually exercised by each of the twelve factors are given in the final column of Table 8.9. It is doubtful whether the management of this company would have agreed upon this weighting of items if the matter had been discussed when the chart was constructed.

The simple adding of ratings for several traits not only fails to weight the traits equally (except occasionally and by chance) but also fails to give them any preassigned weights that might have been decided upon and crystallized in the form of a maximum value that each trait may receive. Suppose, for example, that management has decided that *accuracy* is twice as important as *production* and therefore has adopted a chart in which *accuracy* is rated on a 40-point scale and *production* on a 20-point scale. This arrangement will not necessarily result in *accuracy* being weighted twice as heavily as *production, for the relative weights of the traits are determined by the variability or spread of each and not by the maximum values assigned to each.* It would be quite possible in the situation described above for the production ratings to assume the heavier weights if the original ratings were directly added.

A second hazard related to the matter of weighting items on a merit-rating chart, even when proper steps have been taken to insure the functional operation of the weights decided upon, deals with the weighting of traits for employees on different jobs. Consider, for example, the case of rating two clerical workers, one a receptionist secretary and one a private secretary to the plant superintendent. In many organizations the receptionist secretary, in order to do her job efficiently and satisfactorily, needs a maximum of appearance, poise, tact, and friendliness, and very little, if anything at all, of such traits as originality, versatility and intelligence. The private secretary, on the other hand, needs all of these latter traits, although she perhaps does not need to be quite so personable on first appearance as does the receptionist secretary. Several traits might well operate to give a total over-all rating to the receptionist secretary considerably lower than that of the private secretary; and yet this receptionist secretary might possess the one or two characteristics that such a job requires, so that not only is she highly satisfactory but actually she is able to do this job even better than the private secretary would be able to do it if she were transferred to it. In this connection Locke[31] has suggested that the merit rating system should be keyed to the jobs concerned.

The use of a single over-all trait called "job performance" as a criterion as recommended on p. 227, largely eliminates this problem. A supervisor in rating an individual in terms of his or her ability to do his present job tends to take all of these factors into consideration in an unconscious general evaluation. In the situation referred to above, it is quite likely that the receptionist secretary would be rated in terms of her present job as highly as the private secretary is rated in terms of hers. But if the merit-rating chart is so set up that it is necessary to rate each individual on a number of specific points and then to determine the total merit rating from some combination of these several ratings, it is quite likely that an over-all result not compatible with the facts will be obtained. Here, then, we are presented with a situation in which we must be careful about judging an employee's fitness on his present job by adding the results of several merit-rating items. From this point of view, the halo effect, as Bingham[32] has pointed out, may be the saving grace of a fractional merit-rating system. In other words, when a supervisor begins to rate any given employee on several traits, he may, and probably does, first of all center his attention upon the particular trait or traits that are necessary for the employee on his particular job. After these traits have been rated, all of the remaining traits on the chart automatically fall in line in terms of where the key traits have been rated. But the old axiom that one cannot eat his cake and have it too prevails in the use of an itemized *vs.* an over-all merit-rating system. To the extent that the merit-rating system really does fractionate an employee's characteristics, it is unlikely to give an accurate indication of his ability on his present job; whereas to the extent that the system fails to fractionate the traits, owing to the halo effect, it is likely to work well in terms of giving an adequate over-all indication. Some companies have solved this problem by using a dual system—that is, one in which the employee is first rated on over-all ability on his present job before ratings of a fractionated sort are obtained. Such a dual system will accomplish both results with little more labor on the part of supervisors than is included in any of the systems now in use.

---

[31] N. Locke, "Employee Ratings," *Personnel Journal*, 1943, **21**, 282–288.

[32] W. V. Bingham, "Halo, Invalid and Valid," *Journal of Applied Psychology*, 1939, **23**, 221–228.

RELIABILITY AND VALIDITY OF MERIT RATINGS

The concepts of *reliability* and *validity*, as previously discussed in connection with tests, have corresponding meanings in relationship to merit ratings. The reliability of ratings is the consistency with which the ratings are made, either by different raters, or by one rater at different times. The validity of ratings is the degree to which they are truly indicative of the intrinsic "merit" of employees.

Reliability of Ratings. It can readily be understood that ratings must have an adequate degree of reliability in order to be useful for their intended purposes. For example, if on one occasion the rating for an individual is high and on another occasion it is low—when the individual himself has not changed—one cannot then accept any given rating with confidence. The implication of this statement is that the more unreliable the ratings are, the greater must be the differences in ratings (such as between different employees, or for the same employee at different times) before we are justified in assuming that the differences in ratings reflect *true* differences in merit.

The reliability of ratings is to some degree a function of the type of rating method used. For example, the reliability of rating scales is typically somewhat lower than for employee comparison systems. In connection with the reliability of rating scales, a study by Reymert and Kohn[33] revealed the reliability for a nine-item scale to be .59 when two judges were used. This figure is rather typical of the reliability that may be expected of any merit-rating system based on two or three competent raters.

Data on the reliability of ratings on each of a twelve-item rating chart, as well as the reliability of the total rating, are given in Table 8.10.

The reliability coefficients shown in Table 8.10 were obtained by correlating results from pairs of raters who had rated the same employees. The data are based on a total of 92 raters and 4,500 rated employees.

Somewhat higher reliabilities typically are obtained with em-

---

[33] M. L. Reymert and H. A. Kohn, "The Mooseheart Graphic Rating Scale for Housemothers and Housefathers," *Journal of Applied Psychology*, 1938, **22**, 288–294.

TABLE 8.10

Reliability of Each Item of a Twelve-Item Merit-Rating Scale, and Total Rating Based on the Sum of All Twelve Items

Trait	*Reliability*
1. Safety	.35
2. Knowledge of job	.46
3. Versatility	.47
4. Accuracy	.45
5. Productivity	.46
6. Over-all job performance	.46
7. Industriousness	.47
8. Initiative	.48
9. Judgment	.45
10. Co-operation	.37
11. Personality	.39
12. Health	.36
Total	.55

ployee comparison systems. With a ranking procedure, for example, ratings with reliabilities as high as .85 to .95 have been obtained.[34] Experience with the paired comparison method has indicated that it frequently results in relatively acceptable reliability. In one study, for example,[35] eight men were rated in common by supervisors A and B, eight others were rated in common by supervisors A and C, and eight others were rated in common by supervisors B and C. The results of this study are summarized in Table 8.11. The average correlation of .83 obtained in this study is typical of the reliability that has been achieved when this rating system is used.

TABLE 8.11

Rank-Order Correlations Between Paired Comparison Ratings of Eight Men by Two Different Supervisors

Raters	*No. of men*	*Rank Order Correlation**
A and B	8	.81
A and C	8	.83
B and C	8	.86
Average		.83

* See Appendix A, p. 536, for a description of the rank-order coefficient of correlation.

---

[34] Personal communication from Dr. H. C. Taylor of the Upjohn Foundation for Community Research.

[35] Lawshe, Kephart, and McCormick, *op. cit.*

All of these facts considered together point up the importance of knowing the reliability of ratings obtained with a merit rating system. This statement is not intended to imply that the ratings are of no value, but it does mean that one should be aware of their reliability and should not attempt to use merit ratings in a way that presupposes a higher reliability than they actually possess.

In proportion as ratings are unreliable it is not a valid procedure to consider a slight change in rating from one time to another as indicating a real change in the merit of the employee. By means of a simple statistical procedure, it is possible to obtain for merit ratings what is known as the "standard error of measurement." Unless an employee's rating changes by at least three standard errors of measurement from one rating to another, it is unsafe to assume that any real change has occurred. It is therefore very important for management to know the reliability of the ratings that are used. Only through such knowledge is it possible to know whether the ratings indicate a real difference between two employees or a real change in an employee from one time to another.

Validity of Ratings. Quantitative evidence of the validity of ratings usually is difficult, if not impossible, to obtain. In fact, merit ratings sometimes are used *because* there are no other criteria of employee merit available. Under such circumstances the *validity* of merit ratings may have to be *inferred* from their *reliability*. By this we simply mean that if there is a high degree of consistency between and among raters in rating the same employees, it may be assumed that, *because* of this high reliability, the ratings presumably are indicative of true relative merit (validity). Although high reliability may be used as a basis for inferring the validity of ratings, it does not necessarily follow that low inter-rater reliability indicates that all of the ratings are of low validity. It is possible that low correlation between and among raters may simply be the result of the fact that some of the raters do not know the employees well enough to rate them adequately.

Combining Ratings from Different Raters. It is sometimes the practice to have two or more raters rate each employee, and then combine, or pool, their ratings to obtain a single, pooled rating for each employee. There is evidence to support the contention that pooled ratings *by competent raters* are better than single ratings. This was indicated, for example, in a study by Bayroff, Haggerty, and Rundquist.[36] In this investigation 400 Army officers (primarily majors and lieutenant colonels, attending a Command and General Staff College) were rated experimentally by various methods. They were also ranked by about 20 of their fellow-officers on "value to the Army." These rankings served as a criterion for "validating" the ratings that were made separately. The validity of an eight-step rating scale was determined by correlating ratings from this scale with the criterion rankings. The average correlation for individual raters was about .52 (.52 for one group of 200, and .53 for another group of 200). When combining the ratings of ten raters, however, the correlations were .89 and .84. Thus, if there are several *competent* raters who can rate each employee (as there were in this case) it would seem worth the time and trouble to obtain such multiple ratings.

There is, however, a big "if" in this statement, for it frequently is the case in industry that the only person who really can properly rate an employee is his immediate supervisor. In fact, the very organization of a modern industrial establishment is such that the more ratings one attempts to obtain on a given employee, the further away from that employee in terms of contact with him one must go in order to obtain the additional raters.

The reduction in the validity of merit ratings that occurs when ratings are made or reviewed by supervisors removed from the men at work is illustrated in a study reported by Bittner.[37] Bittner found that although review and endorsement of ratings by the officer at the next higher level of command operated, on the average, to increase the validity of the ratings, there was a *loss* in validity when the endorsing officers were far removed organizationally from the raters. Unfortunately it is not necessary to go very far before one has brought in raters who are so far away from the actual observation of the employee on his job that their ratings of him are not only relatively worthless in themselves, but, what is worse, are also so inaccurate that when averaged with the ratings of the foreman in direct contact with the worker the result is to decrease whatever validity the foreman's ratings may possess. One rotten apple may spoil a bushel of good ones, and it is even more likely to spoil one or two good ones if those one or two happen to be in close contact with the spoiled one. Thus there is considerable danger in assuming that pooling will automatically increase the validity of merit ratings. This result will not be accomplished if the pooling necessitates pooling unreliable with reliable ratings. Thus, it is desirable that a merit rating system provide for the rater to state how well he is acquainted with the employee being rated, and under what circumstances he has had the opportunity to judge the employee. This makes it possible to pool ratings when such pooling is likely to increase their value, and to avoid the pooling if by doing so the ratings would be less valid.

---

[36] A. G. Bayroff, Helen R. Haggerty, and E. A. Rundquist, "Validity of Ratings as Related to Rating Techniques and Conditions," *Personnel Psychology*, 1954, 7, 93–113.

[37] R. H. Bittner, "Developing an Industrial Merit Rating Procedure," *Personnel Psychology*, 1948, 1, 403–432.

Effect of Rater's Memory on Ratings. A condition that often tends to make the ratings more consistent from one year to another than the facts of the situation justify is the supervisor's memory of his previous ratings of the employees. It is only reasonable to expect that when a supervisor rates a man twice, the second rating will be similar to the first if no marked new factors have arisen during the intervening period that might cause the supervisor to change his estimate of the man. The extent to which this memory element influences the consistency of ratings has been studied by determining the correlation between ratings on successive years for groups of employees who were rated by the same and different raters on the two occasions. The findings of this analysis are summarized in Table 8.12.

TABLE 8.12

Correlations Between Ratings on Successive Years of Employees with Same and Different Raters

	Correlation
Same three raters both years	.65
One new rater second year	.59
Two new raters second year	.46
Three (all) new raters second year	.49

An inspection of this table reveals a decreasing trend in the size of the correlation from the situation at the top of the list, where the same three raters were involved during the two successive years, to the bottom of the list, where two or three new raters were involved the second year. These correlations should not, of course, be interpreted as reliability coefficients because the ratings were separated by an interval of a year, and during a period of that length many employees might actually change in such a way as to justify a change in the rating given them. The fact that a rather marked difference exists between the situation in which the same raters are involved and that in which one or more raters have been changed, however, indicates that the constancy of the raters does affect the consistency of the ratings.

A MERIT-RATING CASE STUDY

The merit-rating program of the Acme Steel Company, which has been in operation for several years, includes safeguards against most of the extraneous influences discussed in this chapter. A chart which includes ten factors is used. The factors are:

1. Accuracy of work.
2. Quantity of work.
3. Use of working time.
4. Ability to work with others.
5. Ability to learn.
6. Safety.
7. Initiative and acceptance of responsibility.
8. Conduct on the job.
9. Care and use of equipment, tools, material, supplies, power.
10. Attendance and punctuality.

Each man is rated independently by two or more supervisors who are familiar with the man's work. Each rating is then converted to a common base to eliminate the raters' tendencies. The average corrected ratings by the two or more raters for each man are then determined. The men on a given job or group of jobs in the standard promotional sequence are then listed in order of their corrected average rating. The men in approximately the highest 10 per cent on this list are considered to be equal from the standpoint of the merit ratings and the man in this group with the longest applicable seniority is given the option of accepting any promotion that may be made in the particular promotional sequence involved.

SUGGESTIONS FOR IMPROVING A MERIT-RATING PROGRAM

Many of the facts discussed in the preceding pages of this chapter may be put to practical use by deriving from them a number of suggestions for installing and operating a merit-rating system. Some of the more important suggestions are summarized below:

1. Train the raters.[38] Rating people is a skill, and one that is not easy to learn. Good production men and even good supervisors are not necessarily good raters *unless they have been taught how to rate.* Even at best, rating is subjective and personal. Prejudice and bias can never be completely removed, and in the absence of training they may completely distort many ratings. Companies that have had all raters attend a systematic training program of from six to eight hours before making any ratings have found, without exception, that the time and money spent on this training was a good investment.

Subjects covered in the training of raters include:

a) The "halo effect"—what it is, how it affects ratings, and how to avoid or reduce it.

b) The importance of basing ratings, whenever possible, on facts rather than on opinion. In rating a man's production, specific instances when he failed to meet expected production, for example, are always more convincing than the simple statement "production below par." Very often, specific instances can also be cited in support of ratings on co-operation, judgment, quality, and so on.

c) The necessity for the rater to make up his own mind and to refuse to be influenced by others. When a man has been transferred into a new department, it is not uncommon for the reaction of his former supervisor to his work to accompany him, and so, perhaps, to prevent him from getting off to a fair, fresh start. If his new supervisor rates him on the basis of his former supervisor's judgment, it is possible or even probable that the rating will not adequately reflect his performance on his new job.

d) The importance of using the whole spread on each item or trait. If a forced distribution system is used, the meaning of the guidepost percentages should, of course, be thoroughly covered in the training. But even if a chart system is used, it is important for the raters to realize that the whole spread of the traits should be used if the system is to operate effectively.

---

[38] For a more extensive discussion of the importance of training the raters, see R. S. Driver, "Training as a Means of Improving Employee Performance Rating," *Personnel*, 1942, **18**, 364–370.

2. In addition to training as such, another procedure that increases the accuracy of ratings is to have them made in conference or under supervision. Adoption of this system does not mean that supervisors confer with each other about their men while the ratings are being made, but only that the ratings are made while the raters are gathered for this purpose under the guidance of someone thoroughly familiar with the system. By the use of this procedure, several difficulties will be avoided. Careful thought will be given to the problem, and the hasty checking that often takes place if the supervisor is expected to "find the time" to make the ratings will be eliminated.

3. After the ratings have been made, be extremely cautious in comparing the ratings of men in different departments or on different jobs. The need for this caution is brought out on p. 230. Unless a correction has been made for a man's job, it is unfair to compare his rating with ratings of men on different jobs.

4. Interpret the ratings in terms of actual job demands. A man should not be penalized because a rating shows him to be weak in some ability that his job does not require.

5. Avoid the use of numerical values in discussing ratings with employees. Numerical values give to ratings an appearance of greater accuracy than they in reality possess. The reason for this caution is discussed on p. 234.

6. It is ordinarily wise to omit from the chart factors concerning which objective information can be obtained without using ratings. If production records are available, it is better to use these in evaluating the production of an employee than to rely on a supervisor's judgment. The same thing is true of quality of work, if an actual record of rejects or waste material is available, and of health. A supervisor can judge whether a man *looks* healthy or *acts* healthy, but it is the job of a physician to determine whether he really *is* healthy.

7. Avoid pooling poor ratings with good ones. The use of *one* satisfactory rating is better than the pooling of many, if some of those pooled are likely to be in error. This subject is discussed on p. 245.

8. A final suggestion relative to the most effective use of merit ratings based on rating scales, or indeed of merit ratings based on any system, is to make use of all possible supplementary information

to be obtained from other sources. A very great deal of information about the performance of employees often remains unassembled and unused in many plants. Information on accidents and hospital visits, absenteeism, training time and costs, production, quality, and so forth, frequently is available, and can be very helpful in interpreting or supplementing merit ratings. Job-information tests are also being found useful by a number of plants as one of the tools to be used in upgrading and promoting. In some companies the union has endorsed the use of tests for this purpose when merit ratings alone were not accepted. In a grievance hearing, an unsupported merit rating presented to an arbitrator as reason for an alleged discrimination against an employee with seniority is usually very difficult to justify. But a merit rating supported by one or more test scores on job-information tests or by records covering production, quality, and so forth, often forms a persuasive and conclusive case.

Merit ratings serve their purpose most effectively when used in conjunction with all available additional information about the employee's performance on the job. Merit ratings used in conjunction with such additional information and in the light of possible sources of difficulty discussed in the preceding pages furnish a valuable aid in the administration of an industrial relations program.

9

TRAINING OF INDUSTRIAL EMPLOYEES

Employees must be systematically trained if they are to do their jobs well. No matter how carefully men have been selected, or how much aptitude they may have for their jobs, systematic training is essential if they are to reach a satisfactory level of job performance. Formerly it was standard practice to have training done by the foreman or supervisor in charge of the work. Inexperienced men were sent to the supervisor (or hired by him) and it was his job to see that the new men were given the necessary training. But just as modern scientific management has found that experts should supplement the work of the foreman in making job analyses and setting rates, so also management is finding that the use of experts in training is advantageous as a supplement to the work of the foreman. This does not mean that outside experts need necessarily *do* the training. Their function, rather, is to set up the machinery for the training to be done by properly qualified persons, to supervise the construction or writing of necessary manuals or other materials, and to keep in close touch with the accomplishment of the trainees during the instructional period.

The nature and purposes of training programs vary a great deal, but the purposes generally can be grouped into the following categories:

1. Developing skills (the abilities of employees to perform their work effectively).
2. Transmitting information (such as information about the company, its products or services, its policies).

3. Modifying attitudes (developing more favorable attitudes on the part of employees).

DETERMINING TRAINING NEEDS

Although training programs, generally, are intended to serve the purposes mentioned above, any particular program should be designed to meet certain specified needs. Training for the sake of training is obviously a wasteful activity. Once one has identified the specific training needs, a training program should be developed to achieve these objectives. Where feasible, it is also desirable to measure the effectiveness of the training in achieving the needs originally identified. Thus, before any training is carried out, it is in order to determine what specific purposes the training should fulfill.

The determination of training needs may be clear-cut in some circumstances, but in other situations is far from being a straightforward process. It is probably in part because of the difficulties of determining specific needs that some training programs have been ill-conceived.

The methods of determining training needs range from those which are entirely subjective to those which are quite objective. The nature of some training situations is such that the determination of the needs must of necessity be based on subjective judgments. A survey by Mahler and Monroe[1] dealt in part with the analysis of training needs of 150 companies. In the portion of the survey relating to methods of determining training needs, the training directors of the companies were asked to indicate their preferences for various methods for production, clerical, and supervisory and technical employees. A summary of these expressed preferences is given in Table 9.1. Although there are variations for the three types of employees, it can be seen that, generally speaking, the most common methods preferred were (a) "informal observation," (b) requests from management, (c) talks with supervisors, (d) group discussions, and (e) analysis of reports.

It is evident from the preferences for methods listed in Table 9.1 that the determination of training needs at present is predominantly based on subjective judgments or observations. Although we some-

---

[1] W. R. Mahler and W. H. Monroe, *How Industry Determines the Need for and Effectiveness of Training*, Personnel Research Section, Department of the Army, PRS Report No. 929, 15 March 1952.

TABLE 9.1

Per Cent Preferences of 150 Training Directors for Various Methods of Determining Training Needs

Method of Determining Training Needs	*Type of Employees* *Productive*	*Clerical*	*Technical-Supervisory*
a) Informal observation	12	12	11
b) Requests from management	17	16	15
c) Talks with supervisors	16	17	14
d) Talks with non-supervisors	4	6	3
e) Group discussions	10	10	13
f) Training advisory committees	5	5	6
g) Questionnaires to supervisors	2	2	2
h) Questionnaires to trainees	1	2	4
i) Supervisory morale survey	1	1	4
j) Employee morale survey	4	4	2
k) Tests	3	4	2
l) Merit ratings	4	6	4
m) Intensive interview with supervisors	2	4	4
n) Interview with union officials	2	1	4
o) Analysis of reports (costs, turnover, grievances, etc.)	12	7	7
p) Other	4	3	3
Total	100%	100%	100%

Source: Mahler and Monroe, *op. cit.*

times are inclined to discount the validity of human judgments or observations (frequently with good cause), it should be kept in mind that sometimes the *method* of obtaining judgments or observations affects their validity. In our present context of determining training needs, for example, there have been at least a few training circumstances in which some interesting techniques have been developed.

Check Lists of Training Needs. Two illustrative activity check lists for recording "observations" have been presented by Fryer, Feinberg, and Zalkind.[2] These check lists are shown in Tables 9.2 and 9.3. Table 9.2 is part of a check list to be used by a training specialist in observing the work of a production department. It provides for indicating by a "yes" or "no" check whether or not each check list statement has been observed in the department. As filled out for a hypothetical situation, it can be seen that certain observations suggest possible training needs. The other check list (Table 9.3) is one that would be applicable for identifying, in a similar manner, the possible training needs for employees on a particular type of job, in this specific illustration the job of supervisor.

---

[2] D. H. Fryer, M. R. Feinberg, and S. S. Zalkind, *Developing People in Industry* (New York, Harper and Brothers, 1956).

Judgments of Training Needs. A systematic approach to obtaining judgments of employees regarding training needs of management personnel was developed by Wood.[3] The study will be described in some detail—more to illustrate the method than the results as such.

TABLE 9.2

Check List for Needed Training

Items Recorded by Training Specialist	*Is Item Seen in This Production Unit?*		*Possible Training Need*
	YES	NO	
Downward communications to supervisor slow	x		x
Frequent gripes by supervisors about tools		x	
Supervisors use suggestions from workers		x	x
Many rejects returned to production		x	
Customers' complaints from sales ignored		x	
Turnover in shop higher than seems necessary	x		x
Many lost-time accidents		x	
Paper work up to date		x	x
Raw-material delays		x	
Supervisors use specialists of company		x	x

Source: Fryer, Feinberg, and Zalkind, *op. cit.*

To begin with, about 500 management personnel were asked to list or describe their own training needs, or those of their bosses. These listings were used to develop three questionnaires, each with 68 items. The questions in each questionnaire were similar, except for the frame of reference of the three questionnaires, as follows:

1. Your personal need.
2. Your superior's need.
3. Your subordinate's need.

---

[3] W. F. Wood, "Identification of Management Training Needs" (Ph.D. thesis, Purdue University, June 1956).

TABLE 9.3

Sample Check List for Supervisors

Items Recorded by Training Specialist	*Checked for Adequate Performance*		*Possible Training Need*
	YES	NO	
Keeps inventory of tools	x		
Prepares training outline for apprentices		x	x
Takes unsafe machinery out of service	x		
Checks all repairs	x		
Maintains "hours of work" record	x		
Inspects regularly for quality of product		x	x
Informs on elimination of waste		x	x
Plans workplace layout		x	x
Instructs on cost of materials		x	x
Explains company policy to workers		x	x

Source: Fryer, Feinberg, and Zalkind, *op. cit.*

The items in the questionnaires dealt with seven major areas of training needs as follows (one item from each area is given to illustrate the types of items on the questionnaire):

Area	*Example of item*
1. Subordinate relations.	Coaching my subordinates to help them improve their performance.
2. Personal development.	Getting my ideas down on paper.
3. Labor relations.	Handling a grievance.
4. Wage and salary.	Understanding and applying the company's wage payment policy.
5. Management control.	Understanding the application of departmental budgets.
6. Technical knowledge.	Further technical information on manufacturing processes.
7. Knowledge of company.	How the company products are sold to the consumer.

Each item could be checked by indicating whether an individual felt that he (or his immediate supervisor or subordinate) had "No need," "Some need," or "Real need" for training as indicated by the item.

The statistical analysis of a preliminary try-out with the questionnaires indicated fairly satisfactory reliabilities of the responses for the seven areas (with six of the seven reliability coefficients being .86 or above). The final forms of the questionnaires were administered to management personnel in three plants of a large company, as follows:

Form	*No.*	*Position*
Your personal need	310	Gen. foremen and superintendents
Your superior's need	824	Foremen and gen. foremen
Your subordinate's need	89	Superintendents and top mgmt.
Total	1,223	

One type of analysis made with the resulting data dealt with the degree to which training needs were perceived by *individuals* (self perceptions) as contrasted with their needs as perceived by their *superiors* and by their *subordinates*. Fig. 9.1 illustrates the pattern that emerged in all seven training-needs areas, this particular example being that of management control. In this, and in all the other areas, the *superiors* perceived the training needs of individuals whom they supervised as being greater than did the individuals themselves. In turn, the *subordinates* of the individuals perceived the training needs of those individuals (their superiors) as being *less* than did the individuals themselves.

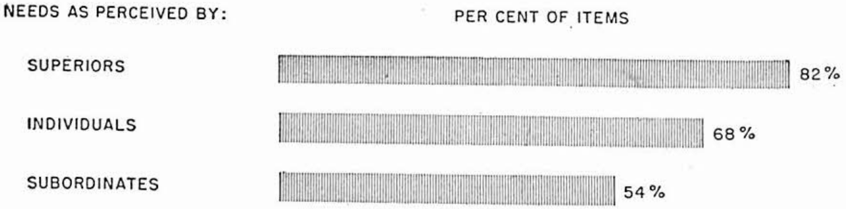
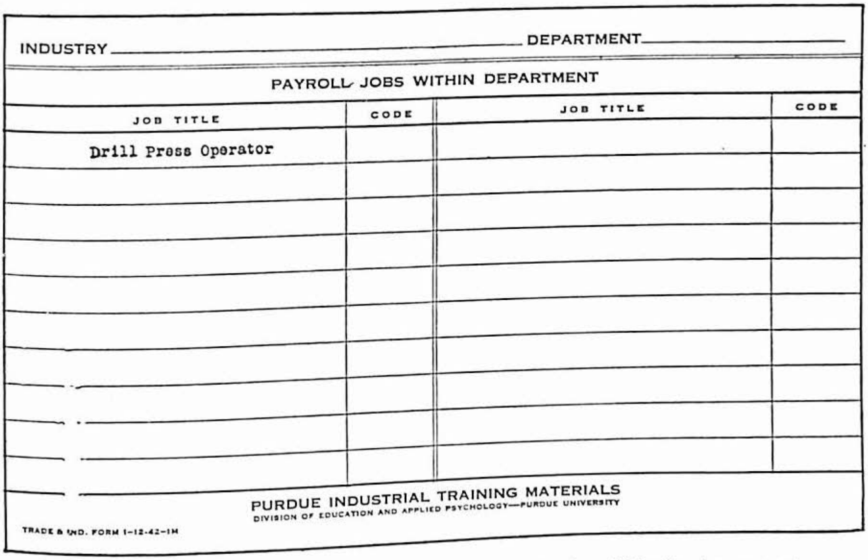


FIG. 9.1. Per cent of training needs items in "management control" area for which need was indicated by individuals, by their superiors, and by their subordinates. (From Wood, op. cit.)

Job Analysis for Training Purposes. A study of training needs identifies the end results toward which training should be directed. Where the training is directed toward the development of job skills, a job analysis of the job in question is the next step in the development of the training program. This involves a very detailed type of job analysis, one that is specifically designed for training purposes.

An illustration of materials that may be used effectively in making a job analysis for training purposes is given in Figs. 9.2, 9.3, and 9.4.



INDUSTRY ____ DEPARTMENT ____

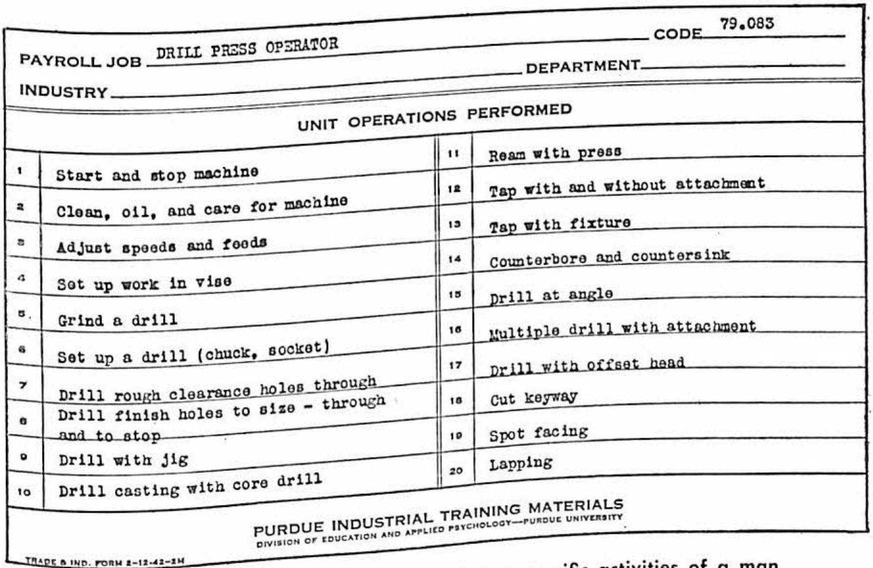
PAYROLL JOBS WITHIN DEPARTMENT

JOB TITLE	CODE	JOB TITLE	CODE
Drill Press Operator			

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY

TRADE & IND. FORM 1-12-42-1M

FIG. 9.2. This is an A Card, used for listing payroll jobs within the department. (Courtesy C. H. Lawshe, Purdue University.)



PAYROLL JOB DRILL PRESS OPERATOR CODE 79.083

INDUSTRY ____ DEPARTMENT ____

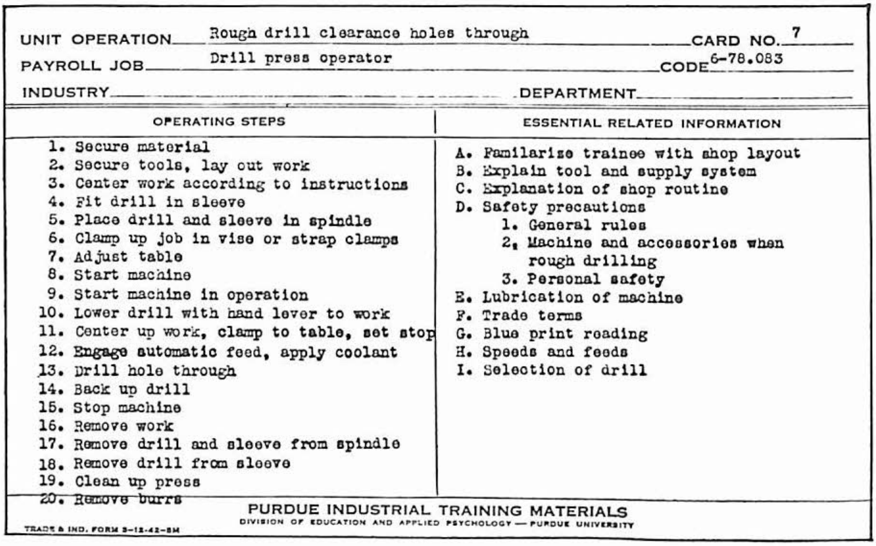
UNIT OPERATIONS PERFORMED

1	Start and stop machine	11	Ream with press
2	Clean, oil, and care for machine	12	Tap with and without attachment
3	Adjust speeds and feeds	13	Tap with fixture
4	Set up work in vise	14	Counterbore and countersink
5	Grind a drill	15	Drill at angle
6	Set up a drill (chuck, socket)	16	Multiple drill with attachment
7	Drill rough clearance holes through	17	Drill with offset head
8	Drill finish holes to size - through and to stop	18	Cut keyway
9	Drill with jig	19	Spot facing
10	Drill casting with core drill	20	Lapping

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY—PURDUE UNIVERSITY

TRADE & IND. FORM 2-12-42-2M

FIG. 9.3. This is a B Card, used for listing specific activities of a man on the job. (Courtesy C. H. Lawshe, Purdue University.)



UNIT OPERATION Rough drill clearance holes through CARD NO. 7
PAYROLL JOB Drill press operator CODE 6-78.083
INDUSTRY DEPARTMENT

OPERATING STEPS

1. Secure material
2. Secure tools, lay out work
3. Center work according to instructions
4. Fit drill in sleeve
5. Place drill and sleeve in spindle
6. Clamp up job in vise or strap clamps
7. Adjust table
8. Start machine
9. Start machine in operation
10. Lower drill with hand lever to work
11. Center up work, clamp to table, set stop
12. Engage automatic feed, apply coolant
13. Drill hole through
14. Back up drill
15. Stop machine
16. Remove work
17. Remove drill and sleeve from spindle
18. Remove drill from sleeve
19. Clean up press
20. Remove burrs

ESSENTIAL RELATED INFORMATION

A. Familarize trainee with shop layout
B. Explain tool and supply system
C. Explanation of shop routine
D. Safety precautions
    1. General rules
    2. Machine and accessories when rough drilling
    3. Personal safety
E. Lubrication of machine
F. Trade terms
G. Blue print reading
H. Speeds and feeds
I. Selection of drill

PURDUE INDUSTRIAL TRAINING MATERIALS
DIVISION OF EDUCATION AND APPLIED PSYCHOLOGY — PURDUE UNIVERSITY
TRADE & IND. FORM 3-12-42-5M

FIG. 9.4. This is a C Card, used for listing the detailed steps and related information necessary for performing each part of the job. (Courtesy C. H. Lawshe, Purdue University.)

Three series of these cards are prepared. On an A card, illustrated in Fig. 9.2, are listed all payroll jobs in the department. For each of these jobs a B Card, illustrated in Fig. 9.3, is prepared. On this card is listed every operation that is performed by a man on this job. For each operation a C Card, illustrated in Fig. 9.4, is prepared. When the analysis is complete for any given job, there will be as many C Cards for that job as there are operations to be performed. The C Cards thus become the training manual for the job. A careful use of this kind of analysis in training new employees avoids many omissions in the training that are almost certain to be made if some systematic procedure is not followed. An application of this method of job analysis for a specific operation is given in Table 9.4.

A unique type of job analysis for training purposes is illustrated by the work of Lindahl.[4] Although the particular method he used was specific to the job in question, this example illustrates how a bit of ingenuity can be useful in developing a method appropriate to a particular circumstance. He set up a simple graphic recording ap-

---

[4] L. G. Lindahl, "Movement Analysis as an Industrial Training Method," *Journal of Applied Psychology,* 1945, **29,** 420–436.

TABLE 9.4[5]

How to Prepare the Strip or Foxing Calender for Operation

Step	*What to Do*	*How to Do It*
1.	Advise mill man of first compound required.	a) Check ticket. b) Indicate type and approximate quantity of run.
2.	Advise take-off crew of type of material being processed.	a) Check list. b) Indicate type and approximate quantity of run. c) Indicate whether boards or rolls will be used.
3.	See that correct engraved roll is in calender.	a) Check list. b) The #4 roll is interchangeable to obtain correct engraved rubber surface. c) See breakdown titled "How to Change a Calender Roll."
4.	Apply heat on all four rolls.	a) Be sure couplings are fastened. b) Open individual valves for each calender roll. c) More steam on rolls #3 and #4. d) Less steam on rolls #1 and #2.
5.	See that all moving parts are free.	a) Be sure static discharges are out of way. b) Be sure no foreign object is near the bite of the calender rolls on the drive gears.
6.	Start machine and test safety devices.	a) Start motor (usually a separate button). b) Push starting button on central panel. c) Decrease speed to slowest speed—as soon as calender starts. d) Pull safety stop cable from both sides of machine. e) Test stop button on control panel.
7.	Apply heat on next roll required.	a) Check ticket. b) Locate next roll on storage rack. c) Apply heat at storage rack.
8.	Place cutting bar on calender (if running strips or foxing).	a) Check ticket. b) See that correct cutting bar is in place on the calender above roll #6. c) See that cutters are in "no cut" position, "dog" released and lever up.
9.	Adjust boards at back of feed rolls #1 and #2.	a) Set slightly narrower than desired width of finished calendered stock.

---

[5] By permission of the Personnel and Training Section of the United States Rubber Company, New York, N.Y.

TABLE 9.4 (Cont.)

Step	*What to Do*	*How to Do It*
10.	See that rolls are clean.	a) Clean surface with gasoline and cloth. b) Clean excess grease and oil from edges of rolls. c) Be sure all rubber scrap removed from crevices at edges of rolls.
11.	Check to see that calender crew is ready for run.	a) Warming mill man has proper stock at correct temperature. b) Take-off crew has correct reels or boards.

paratus to show the stroke pattern followed in cutting tungsten discs. The machine studied made use of a rotating abrasive wheel that is lowered by means of a foot pedal to cut off a disc. After the disc has been cut, the foot pressure is released, the abrasive pressure is relieved, the abrasive wheel rises, the tungsten rod advances, and the machine is ready for a repetition of the cutting cycle. The recording apparatus used by Lindahl is illustrated in Fig. 9.5. Identification of the various parts of the cutting cycle is shown in Fig. 9.6.

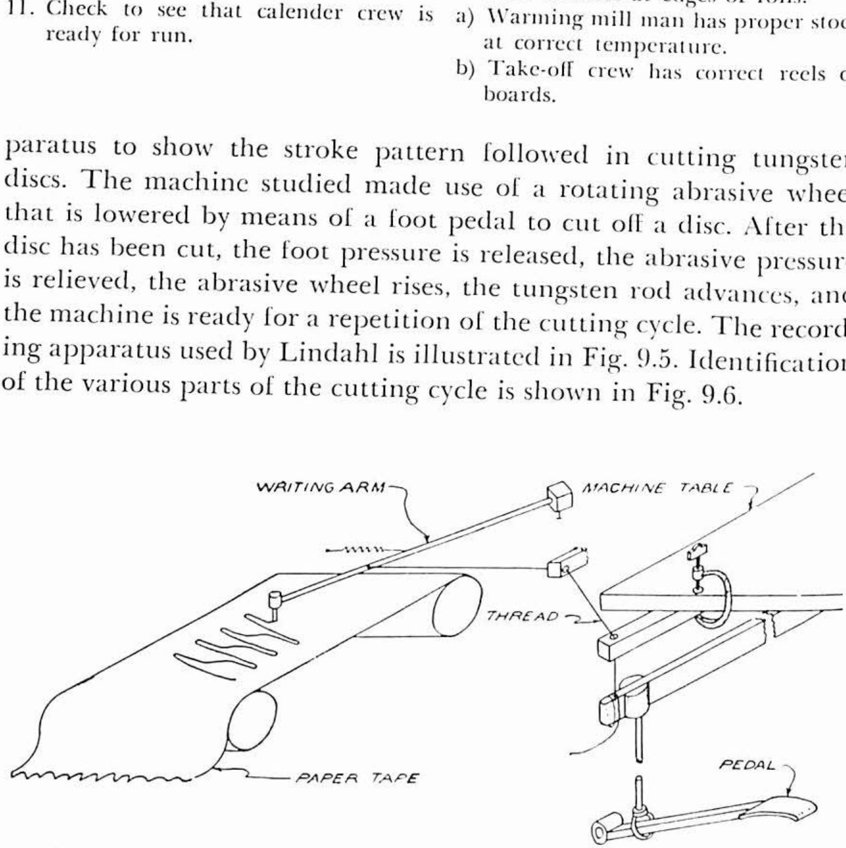


FIG. 9.5. Schematic drawing of apparatus used in obtaining graphic records of job performance in cutting discs with a foot-operated abrasive wheel.

The cutting patterns found at various stages of training in the performance of a typical trainee are illustrated in Fig. 9.7. This particular method of analyzing the job also lent itself to use as a

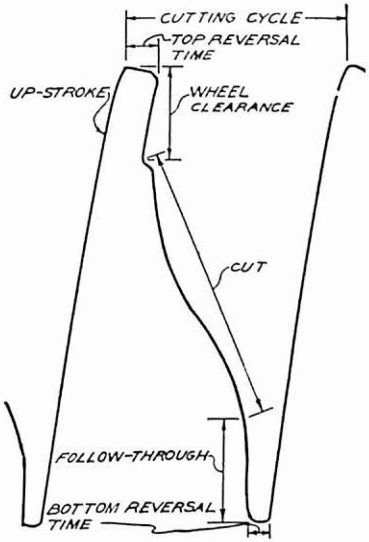


FIG. 9.6. Identification of various parts of the cutting cycle as shown by the graphic record attained with the apparatus shown in Fig. 9.5.

training device. By using this simple equipment to show trainees the specific errors they were making in their cutting cycles, Lindahl was able to improve the quantity and quality of work at various stages of training, and also to increase the life of abrasive wheels.

CHARACTERISTICS OF LEARNING

Human learning has been studied extensively by educational psychologists. Their work has been largely concerned with learning by children in school, but many of their findings and almost all of their basic methods of investigation are directly applicable in the industrial situation, i.e., to employees learning their jobs instead of children learning their school subjects. Two of these general findings deal with the *learning curve* and the *plateau*.

The Learning Curve. In investigations of learning it is frequently the practice to obtain a "learning curve." A learning curve indicates the progress of individuals, or groups of individuals, in the learning

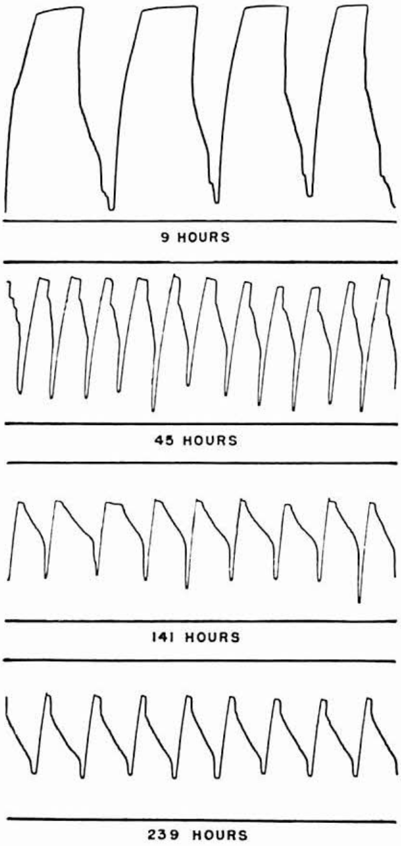


FIG. 9.7. Records of performance of a trainee at various stages of his experience on the job.

situation under investigation. Such a learning curve for the operation of looping in the hosiery industry is reproduced in Fig. 2.5 on p. 21, and a similar curve is reproduced on a card distributed to new employees, as illustrated in Fig. 9.10. Learning curves for the operations of trimming, covering, and hemming in the textile industry are reproduced in Fig. 9.9 on p. 277.

The general learning curve for a given operation shows at a glance the level of skill that has been reached by the average operator after a given period of training or experience on the job. Anyone who is responsible for the training of new employees on a specific job will find that a learning curve obtained from the records of previous employees is of considerable value. By comparing the learning curve of an individual with one based on the average for many individuals, it is possible to see immediately whether the progress of the individual is up to standard and, if not, how far the performance is below standard at any time in the instructional period. He will also be able to detect immediately whether an employee is maintaining his former level in comparison with the standard curve or is gaining or losing in comparison with other employees. In short, he will have a basis of comparison by which he can evaluate the effectiveness of his own instruction and the effectiveness of the learning that is being done by the new employees. If an employee falls below the curve by a significant amount, and, in spite of every effort of the employee or the instructor to remedy the situation, continues to lose ground in comparison with the curve for the average, it is possible that the employee should be transferred to some other job before too much time and effort have been wasted in training him. It is even desirable, in some instances, to supply each new employee with a copy of the general learning curve for the job and allow him to plot his own curve on the same chart, thus giving him a graphic record of his progress from day to day or week to week.

It should not be inferred from the foregoing discussion that there is any generalized form of learning curve that can be reduced to a mathematical equation and that applies to all learning situations. Many factors determine the form of a learning curve, and these factors vary greatly from one situation to another. Since one of these factors is the method of training, it is not uncommon to find different learning curves for the same industrial operation when different methods are involved. Indeed, a comparison of the resultant learning curves for different methods of training is an excellent way to compare the effectiveness of the methods. But with the various factors, such as method of training, held as constant as possible, a reasonably standardized curve results for a given operation for the average person. It is this learning curve that the job trainer should have available if he is to evaluate adequately the effectiveness of his instruction and the progress of the trainees who are under his supervision.

The Plateau. In learning any complex task it often happens that, after a certain level of efficiency has been attained, a period of time arrives in which little or no improvement takes place. This period is followed by a later increase in skill. The period during which no apparent improvement occurs is known as a *plateau*. Plateaus are not always present in the learning curves for industrial operations, and when they are present it is sometimes possible to eliminate them by changed or improved methods of instruction. No matter what method of instruction may be in use, however, arrival at a plateau in the learning process often is characteristic of the average learner's performance.

A typical example of a plateau in a learning curve is shown in Fig. 2.5 on p. 21. We do not need to concern ourselves here with the cause of the plateau, although it might be mentioned that a considerable amount of psychological investigation has been devoted to this subject and a number of experiments have been made that have identified certain factors that may result in a plateau.[6] If a plateau is characteristic of the average learning curve of an industrial operation it is highly important for the job trainer who is supervising this training to be aware of the situation. Without such knowledge he may feel that his training is inadequate, that the trainees have lost interest, that they have reached the maximum level of which they are capable and therefore need no further training, or that many of the employees should be shifted off the job because they are unable to reach the level of production achieved by experienced operators. If, on the other hand, he is aware of the existence of the plateau, he will continue his instruction, encourage the trainees to stay with the job until the end of the plateau is reached, and hold in abeyance his final judgment of the trainees' ability until they have had an opportunity to benefit from the final spurt in production that is almost certain to follow the plateau if the training is continued.

By analyzing the activities occurring during the plateau period it is often possible to modify the training program so that the plateau may be partially or entirely eliminated. Whether or not the job trainer is able to achieve this result, his knowledge of the existence of the plateau, together with his awareness of the fact that this situation is by no means uncommon in the learning of industrial tasks, will enable him to cope with the situation in a more intelligent manner and to help tide the employees over a period that, at best, is likely to involve a certain amount of discouragement.

---

[6] W. H. Batson, "Acquisition of Skill," *Psychological Monographs*, 1916, 21, 1–92.

PRINCIPLES OF LEARNING

The problems of employee training may properly be considered as a branch of educational psychology or the psychology of learning. A great many investigations dealing with the problems of learning have been conducted; from these investigations various general principles, part of which are applicable to the industrial training situation, have emerged. The application of some of these principles in the development of training programs generally will increase the effectiveness of training.

Knowledge of Results. One of the principles of learning that has been established from psychological investigations deals with "knowledge of results." When trainees know how they are doing, their learning typically is increased. This principle was first pointed out in connection with the learning of arithmetic among school children. Panlasigui and Knight[7] found that if children are consistently informed of all errors made in arithmetic they show a significantly more rapid improvement in learning the arithmetic computations involved than if they are given exactly the same sets of drills and are not provided with a "knowledge of results." Book and Norvell[8] found the same principle operating in the learning of mental multiplication. In their experiments the persons in the experimental group watched the record of their progress and attempted at each practice to make a higher score than the score made previously. Persons in the other, or control, group were not informed of their progress. These results are summarized graphically in Fig. 9.8. Curve II in Fig. 9.8 shows the results with the experimental group—the one that *was given knowledge* of results until the 10th practice period and no knowledge of results from the 10th to the 15th practice period. Curve I shows the results with the control group—the one that *was given no knowledge* of results during the first 10 practice periods but was given such knowledge from the 10th to the 15th periods. The curves in Fig. 9.8 show that immediately after the tenth practice the groups changed their relative standing. The group that was then given knowledge of results (group I) forged rapidly ahead of the other group.

---

[7] I. Panlasigui and F. B. Knight, "The Effect of Awareness of Success or Failure," *Twenty-Ninth Yearbook of the National Society for the Study of Education*, Part II (Bloomington, Ill., Public School Publishing Company, 1930) 611–619.

[8] W. F. Book and L. Norvell, "The Will to Learn," *Pedagogical Seminary*, 1922, **29**, 305–362.

One need not look long in industry to find situations parallel to this experiment. The work of Lindahl[9] previously mentioned is an example of the application of the principle of knowledge of results to an industrial learning situation. In the investigation of inspectors summarized in Chapter 13 the accuracy of many inspectors was found to be significantly below that of which they were capable and also below the level of accuracy required for satisfactory job performance. It was found that the unsatisfactory level of accuracy characteristic of many employees in this department was traceable to the training

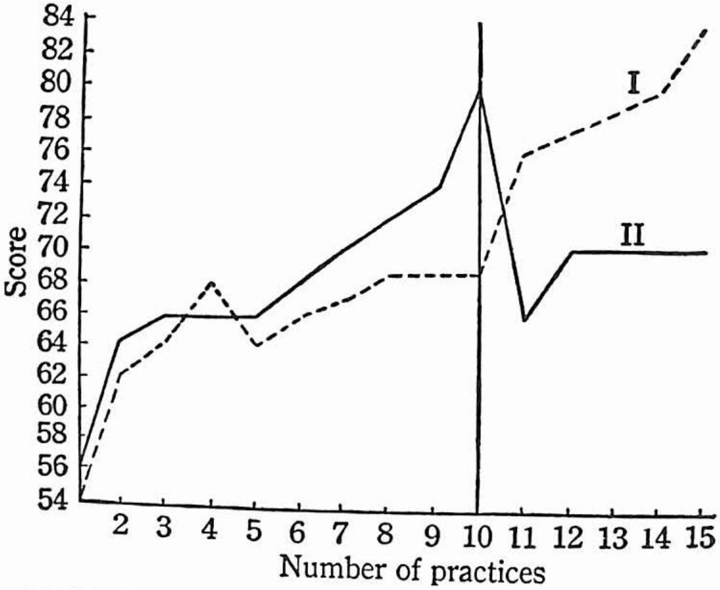


FIG. 9.8. How knowledge of results affects performance in learning a new task. (See p. 265 for discussion.)

---

[9] Lindahl, *op. cit.*

procedure that had been followed prior to this experimental work. The former training method consisted of placing each new inspector between two older inspectors who were asked to instruct the new employee in the details of the job. Since each of the older employees had a definite amount of work to perform during the day it was only natural that they should spend little if any time in training the new employee. Indeed, it was found in many instances that the older employees resented this additional training load and as a result did not give the new employee adequate training in the details of the inspection job. This method of training resulted in a lack of standardization of inspection practice among the employees—many of the inspectors were overlooking certain defects that they were *able* to detect but that they had not been trained to observe. The new training program instituted in this department included the assignment of specific job trainers to the task of teaching new employees, so that the latter would be able to develop satisfactory standards of performance without guessing which sheets were satisfactory and which were defective.

The general principle of job training is that new employees will improve very slowly—if, indeed, they improve at all—unless they are provided with systematic and accurate information on the quality of their work. Provisions for furnishing new employees with definite knowledge of results should be an integral part of any training program.

Motivation. Ordinarily speaking, people will not learn very much about anything unless they are motivated to do so, that is, unless they are supplied with an adequate incentive. Incentives in industry are usually considered under two general heads: financial and nonfinancial. Under financial incentives are considered those that pertain to methods of wage payments, such as an hourly rate vs. some form of piece work. One of the major problems in industrial management is that of deciding upon a wage payment plan that will result in a satisfactory level of job performance and at the same time will be considered by the employees to be fair. If the plan in use does not place a proper premium upon high production, it often happens that the employees never bother to learn the skill required for a very high level of job performance.

A typical example of this situation is furnished in a study by

Kitson.[10] Kitson examined the production records of forty hand compositors whose experience ranged from one to twenty-seven years, the average being eight years. The average output of these employees was 55 units for a week of work on a scale of efficiency established by time studies. An incentive system of wage payment was installed that permitted the employees to earn a bonus for production over the previous average amount. Within five months these employees were averaging 97 units on the scale and at the end of a year the average output had reached 103 units. This experiment made it clear that the previous method of wage payment did not provide sufficient motivation for these employees to learn the skill necessary for high production, even though they had been on the job a number of years.

Incentives are not limited to financial incentives. A number of investigators of this subject, for example Lee[11] and Miles,[12] have stressed the fact that such factors as recognition of good work, a fair system of promotion, and job security all lead to an attitude of mind on the part of the employee that helps develop a desire to learn the job well and to turn in a satisfactory performance on the job.

Another aspect of motivation upon which the evidence is reasonably clear and convincing is the relative desirability of rewards and punishments. Translated into industrial terms, this question raises the issue whether a new employee will learn faster and better if he is complimented for all good job performance or if he is criticized for all poor job performance. Since any new employee will do some things correctly and other things incorrectly, it is quite possible to use either or both of these methods while he is learning the job. Judging from a number of related investigations in the field of experimental psychology, it seems clear that praising satisfactory performance is the more desirable method to follow.[13] Modern industrial practice has largely adopted this point of view, primarily because it results in more satisfactory industrial relations.[14] It is interesting and fortunate that experimental studies in psychology substantiate this judgment for the quite different reason that praise is superior to criticism as a means of inducing the employee to learn.

---

[10] H. D. Kitson, "Extra Incentive Wage Plans from a Psychological Viewpoint." *Bulletin American Management Association, Production Executive Series,* 1925, **9**, 7.

[11] C. A. Lee, "Some Notes on Incentives in Industry," *The Human Factor,* 1932, **6**, 180–186.

[12] G. H. Miles, "Effectiveness of Labour Incentives," *The Human Factor,* 1932, **6**, 53–58.

[13] E. L. Thorndike, *Effects of Punishment and Reward* (Chicago, University of Chicago Press, 1932).

Attention to Specific Features of the Job. All but the simplest of industrial jobs require several different types of skill or ability. The stenographer must be able not only to take shorthand but also to use a typewriter. The inspector is often required to spot certain defects of appearance visually, other defects by weight or feel, and perhaps still other defects by means of readings taken from special instruments. It often happens that the several aspects of a job are sufficiently independent of one another to necessitate giving special training for each. Reference to Table 13.2 on p. 386 shows that the inter-correlations between accuracy of inspectors in detecting different kinds of defects are in general quite low. This means that the inspectors who are best in detecting certain defects are not necessarily better than average in detecting other defects. To learn an inspection job efficiently, new inspectors cannot simply be told to "throw out the defective material." They must be given specific instructions in the nature of the defects and taught what to look for in each specific case. Training should be directed toward the specific elements of the job.

Transfer of Training. Does training on one industrial job transfer to another job? Will an employee who has been thoroughly trained in one operation be easier to train in a new operation than an employee who has not had this previous training? These questions have stimulated a great deal of investigation by psychologists. At one time it was felt (with little or no experimental evidence to support the feeling) that training would develop certain general "powers" or "faculties" so that a person trained in one kind of activity would be more able to do many other kinds of activities in which these faculties might be used. It was felt that the memorizing of Latin verbs and conjugations would improve one's memory so that he would be better able to memorize insurance premium tables or grocery price lists. In industrial terms, training in the operation of one machine, it was felt, would improve an employee in ability to operate some other machine with which he had no direct experience.

---

[14] J. J. Jackson, "Reprimanding Employees," *Personnel Journal*, 1941, **19**, 73–80.

This notion—the wholesale *transfer of training*—has been subjected to a great deal of experimental work, beginning with that of Thorndike and Woodworth.[15] The conclusion reached by these early investigators—and substantiated by several more recent studies—is that the amount of transfer from one skill or mental function to another is far less than was formerly believed. Indeed, Thorndike[16] has concluded that transfer in a general way does not occur at all and that what is often regarded as transfer is simply due to *identical elements* in the two jobs under consideration. These identical elements may be activities (usually manipulative activities in industry), or methods of work. An example of the former would be any routine assembly operation that is identical with the operation in some other assembly job. An example of the latter would be the methods of laying out work that would be similar from one job to another in the activities of a motion study or layout man.

Training for industrial jobs should, except in very rare instances, be specific and to the point rather than haphazard and general. Often a training procedure will have a measurable effect upon one or more specific aspects of a job, and will have no effect whatever on other phases of the operation. For example, Frisby[17] has shown that a landing machine trainer (a training device) speeds up learning in landing and take-off but has no appreciable effect on other aspects of flying.

The conclusion to be kept in mind is that supervisors should not expect more transfer to occur than can be expected in terms of definite similarities in the jobs. Because an employee has been able to learn and successfully perform one job is no guarantee that he will do as well on some unrelated job. On a second job he may succeed or he may fail, but his performance is neither limited nor determined by his performance on the first job.

---

[15] E. L. Thorndike and R. S. Woodworth, "The Influence of Improvement in One Mental Function upon Efficiency of Other Functions," *Psychological Review*, 1901, **8**, 247–261, 384–385, 553–564.

[16] E. L. Thorndike, "Mental Discipline in High School Studies," *Journal of Educational Psychology*, 1924, **15**, 1–22, 83–98.

[17] C. B. Frisby, "Field Research in Flying Training," *Occupational Psychology*, London, 1947, **21**, 24–33.

This principle is recognized by many supervisors and personnel men. The frequent shifting of an employee from job to job until one is found that he can perform satisfactorily indicates that the failure of an employee on one job does not necessarily indicate that he will fail on another.

Negative Transfer or Interference. If the individual who has taught himself the "hunt and peck" system on a typewriter decides to learn the touch system, he must give up the use of all, or most of, the habits he has already formed. It is usually as difficult to break old habits as to learn new ones, and sometimes much more difficult. Thus, a person who has thoroughly learned the wrong way to do a thing before he is taught the right way to do it has the double task of unlearning old methods and learning new ones.

Unlearning errors is a most wasteful form of learning.[18] It is not surprising, therefore, to find experiments showing that when learners have previously been taught *one* method of doing a task, they have more difficulty in learning a new method than do learners who have had no previous experience on the job at all. This principle is known as negative transfer, interference, or inhibition. Its application to the industrial situation suggests that the time to teach employees the correct way to do a job is when they are placed on the job, not after they have had a few days or weeks of experience in learning incorrect methods.

Distribution of Practice. In many learning situations the length and spacing of practice periods is important for efficient learning. Generally speaking, people learn more effectively when practice is split up into periods than when it is scheduled continuously. The optimum lengths of practice periods and of interspersed changes (either rest or other activity) vary with the material to be learned.

A study on the distribution of practice on an industrial job is reported by Mahler and Monroe.[19] The training in question was job instruction training to supervisors, designed to qualify the supervisors to be more effective in giving job training to new employees. One group of 300 supervisors received six hours of training spread over a two-week period. Another group received the six hours in three two-hour sessions on successive days. In a later follow-up it was found that the first group (for which the training was spaced over two weeks) made fewer mistakes in training new workers than did those in the second group (for which training was concentrated in three days).

---

[18] E. L. Thorndike, E. O. Bregman, J. W. Tilton, and E. Woodyard, *Adult Learning* (New York, The Macmillan Company, 1928), 183.

[19] Mahler and Monroe, *op. cit.*

TRAINING METHODS AND AIDS

Among the various methods of training are the following:

1. Lecture.
2. Demonstration.
3. Film.
4. Conference or discussion.
5. Role playing.

The lecture is used in a wide variety of training situations. The demonstration and film methods probably are used predominantly in job-skills training. The conference-or-discussion and role-playing methods are used primarily in supervisory and management training. Although certain methods obviously can be used for the same training purposes, it has been found that for a given purpose one method may be superior to another. An example of an evaluation of training methods will be given, to indicate how different methods can be compared. The study in question is one reported by Fryer[20] on three methods of training Air Force personnel in code learning. Four groups of men, equivalent in initial ability, were covered by the study. The training for the four groups was as follows:

1. Lecture (oral instruction with slides).
2. Manual (students read illustrated manual).
3. Film (15-minute animated film).
4. No training.

The content of the three training methods was the same.

Immediately after training, and again after two months, an objective 25-item test was given to the groups with results as shown in Table 9.5. The results in Table 9.5 show that the film method was superior, both immediately following training and after two months, to either of the other two methods investigated.

---

[20] D. H. Fryer, "Training with Special Reference to its Evaluation," *Personnel Psychology*, 1951, 4, 19–37.

TABLE 9.5

Test Scores for Four Air Force Groups Trained in Code Learning by Different Methods

Method of Training	*Average Test Scores*	
	Following Training	*After Two Months*
Film	17.9	16.3
Manual	15.4	13.0
Lecture	15.2	13.0
No training	5.4	7.0

Source: Fryer, *op. cit.*

Films have also been used successfully in certain types of industrial training. VanderMeer[21] reports a study in which sound films were used in the training of engine lathe operators. He found a significant superiority in the operators trained by this method when compared with those trained by traditional methods.

These studies showed a superiority of the film method over the others. It might be added that films are particularly effective as aids in showing certain mechanical operations, or in presenting material that cannot be demonstrated in a training group. Although films are very effective for some training purposes, they are, of course, inappropriate for others. When used they should supplement the instructor, not supplant him.

Without discussing all of the pros and cons of various training methods, it can be pointed out that for some training purposes "participation" of individuals is a very effective procedure. This is particularly true in supervisory and management training, where group discussion, role playing, and other participation methods are particularly pertinent.

There is general acceptance of the notion that training aids add to the effectiveness of training, and there is experimental evidence to lend support to this belief. There is, however, also a danger in the use of training aids, the danger being that the instructor will become "training-aid happy" and rely on them to the detriment of the effectiveness of the training. In other words they should be used judiciously, and, where used, they should be appropriate to the purpose.

---

[21] A. W. VanderMeer, "The Economy of Time in Industrial Training: An Experimental Study of the Use of Sound Films in the Training of Engine Lathe Operators," *Journal of Educational Psychology*, 1945, **36**, 65–90.

Among the various types of training aids in use are the following:

1. Movies (previously mentioned).
2. Slides and film strips.
3. Charts.
4. Scale models, mock-ups, training devices.
5. Blackboards.
6. Flip-sheets (to supplement a lecture with illustrations that are flipped over on a stand, as needed).
7. Printed materials (books, manuals, hand-outs and so forth).
8. Television.

A technique used by the armed forces with great success during World War II was the tachistoscopic airplane and ship recognition training. This approach, founded upon the basic work of Renshaw,[22] involves presentation of stimuli for a very brief period (often 1/10 of a second or less). By the use of this method, subjects have been taught to see and identify distinguishing features of objects or forms when the "flash" presentations are reduced to an exceedingly small fraction of a second. In fact, Renshaw found that when trained by this method ordinary people rapidly achieve a degree of skill heretofore considered genius. The application of this approach to aircraft recognition training has been described by Terrell.[23] Although this method of training has apparently not been used extensively in industry, it would seem to offer real possibilities for certain types of training, particularly on such jobs as inspection for appearance.

Fryer, Feinberg, and Zalkin[24] suggest the following ten-item check list for use in evaluating training aids. This check list will help the trainer to select the most effective device:

1. *The aid must help the trainee to be more critical.* In the use of films and film strips, their dramatic appeal must not lull the trainee into passive acceptance of the subject-matter. The aid and the manner of presentation should stimulate independent and critical thinking.

---

[22] S. Renshaw, "The Visual Perception and Reproduction of Forms by Tachistoscopic Methods," *Journal of Psychology*, 1945, **20**, 217–232.

[23] C. M. Terrell, "Recognition Training in the Army Air Forces," *Military Review*, Fort Leavenworth, 1945, **24**, 40–43.

[24] Fryer, Feinberg, and Zalkin, *op. cit.*

2. *The aid must give an accurate picture of the concept which it represents.* A training aid must not distort the presentation. The scale must be accurate so as not to confuse or exaggerate the object.

3. *The aid must be suitable to the level of abilities and experience of the trainee.* A training aid must be selected that is neither too advanced nor too simple for the intelligence and educational background of the trainee.

4. *The aid and all its significant parts must be seen clearly by both small and large groups.*

5. *The aid must be easy to maintain and repair.* An aid that breaks down frequently and requires frequent readjustments will be too discouraging to use.

6. *The aid must be economical to ship or store.* Nothing on it should be so small or unique as to encourage pilferage.

7. *The aid should provide for a test of learning progress.* The trainee should be able to recognize the actual article after seeing the aid and distinguish it from among several similar objects.

8. *The aid should be inexpensive, rugged, and safe.*

9. *The aid must be properly designed.* Glare must be controlled, height of mounting should be realistic, and all matters of set-up should be arranged properly before the presentation.

10. *The aid must have "appeal."* It must be attractive in color, movement, or form, but not at the expense of clarity and accuracy. For example, the improper use of a bright color might direct the attention of the trainee away from significant parts.

THE MEASUREMENT OF TRAINING

As implied earlier, training should be carried out in order to fulfill specific training needs. After it is all over, then, it is necessary to determine to what extent the training has accomplished its intended purposes. As pointed out by Mahler and Monroe,[25] the prevailing tendency in evaluating the effectiveness of training is to rely on "opinion." A few companies, however, attempt to measure the effectiveness of training in more objective terms, and there is a trend toward greater emphasis on such measurement.

The measurement of the effectiveness of training obviously requires an appropriate criterion (meaning a relevant and reliable criterion). Many of the types of criteria mentioned in an earlier chapter in connection with test validation also can be used in evaluating the effects of training. Some of these will be discussed below, as they relate to training, with some illustrations of their use.

---

[25] Mahler and Monroe, *op. cit.*

Production. Job-skill training is to train people for production, and therefore the success (or failure) of the training is usually indicated by the production achieved by the trainees at various stages of their training.

In many industrial jobs the work has been time-studied, rates based on these studies have been set, and output or production of employees is indicated by production or earnings. A comparison of production from one time to another reveals immediately how effective the intervening training has been and whether new or different methods should be tried out. A typical example of this method has been reported by Blankenship and Taylor[26] for the operations of trimming, covering, and hemming among textile workers. The learning curves for these three operations, as published by Blankenship and Taylor, are reproduced in Fig. 9.9. In order that these curves may be comparable from one operation to another, the production in each case has been plotted in terms of Bedaux units. These units consider the actual motions involved in each operation and equate these so that a given number of Bedaux units means a given amount of production, regardless of the operation involved. Thus, thirty Bedaux units of production in trimming may be considered equal to thirty Bedaux units in covering or hemming.

It will be noted that for each operation employees with less than five weeks of experience produced from thirty-one to forty-one Bedaux units. With increasing experience, the output in each case rose until, for trimming and hemming, it reached sixty-nine units at the end of thirty-five weeks; while in the case of covering, only sixty-two units were reached in this length of time. It is apparent that no serious discrepancy in speed of learning exists among the three operations studied. It will also be apparent, however, that if a serious discrepancy did exist, that is, if employees on any one of these jobs tended to lag consistently behind employees on the other jobs in speed of learning, an exact quantitative measurement of the amount of the lag would be indicated immediately by learning curves of this type. Such curves, therefore, serve the very useful purpose of providing a means of evaluating the successfulness of an operator-training program and spotting decisively those operations in which training is inadequate, either in quantity or quality.

---

[26] A. B. Blankenship and H. R. Taylor, "Prediction of Vocational Proficiency in Three Machine Operations," *Journal of Applied Psychology*, 1938, **22**, 518–526.

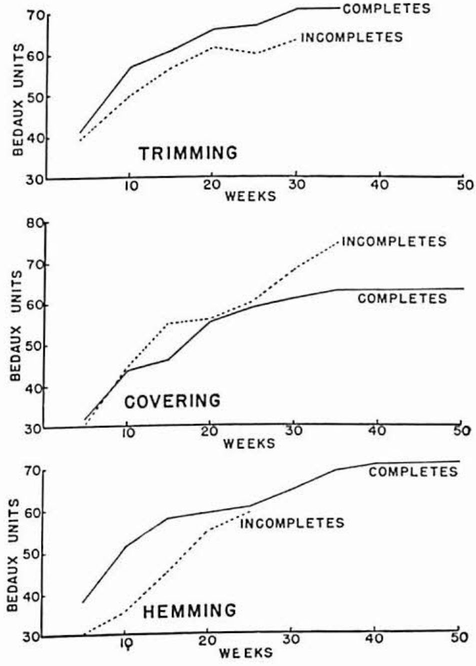


FIG. 9.9. **Learning curves for three operations in a textile mill.**

The question may be raised whether such curves as those in Fig. 9.9 indicate the effect of *training* or the effect of *experience* on the job. The answer depends on the type of activity that has taken place during the period studied. Obviously, if no formal training has been given and each operator has been left free to pick up the "tricks of the trade" in whatever way he or she can, the curves would indicate only the improvement due to experience. On the other hand, if formal training has been given systematically over a period of weeks, the curves will indicate the amount of improvement in job performance occurring as a result of the training and experience. The curves simply show the change in job performance that takes place. Whether this change is due to training or experience or to a combination of these factors can be decided only in the light of a complete knowledge of the conditions under which the operators have been working. But when the curves for a given operation are available, together with a knowledge of the working conditions, it is possible to evaluate in objective and accurate terms the effectiveness of the method that has been used in training new operators.

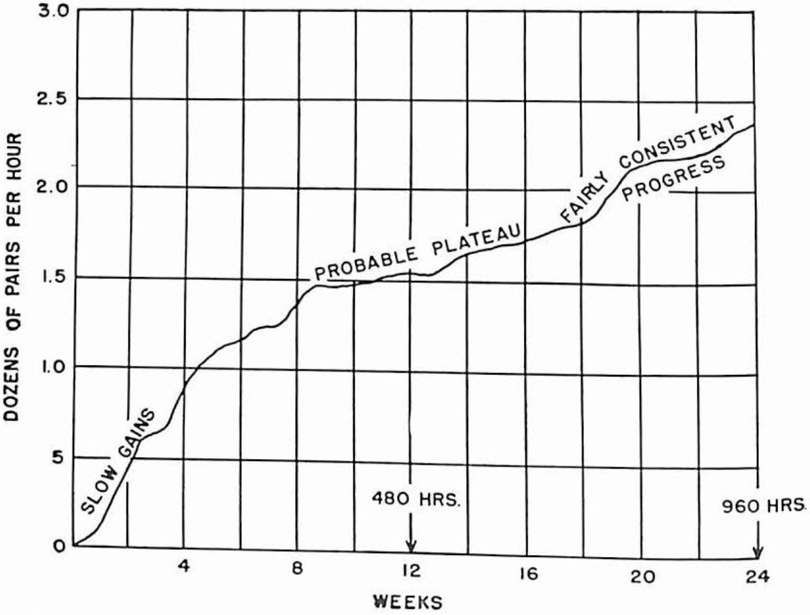


FIG. 9.10. Learning curve of hosiery loopers. (Courtesy J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N.C.)

A similar curve showing the relationship between production and experience on the job of looping in a hosiery mill has been reproduced on p. 21 in connection with the discussion of individual differences among operators at different stages of the learning process. Another curve of learning for hosiery looping from another plant is reproduced in Fig. 9.10, a copy of which is given to all new employees on this particular job in this plant. Although these curves

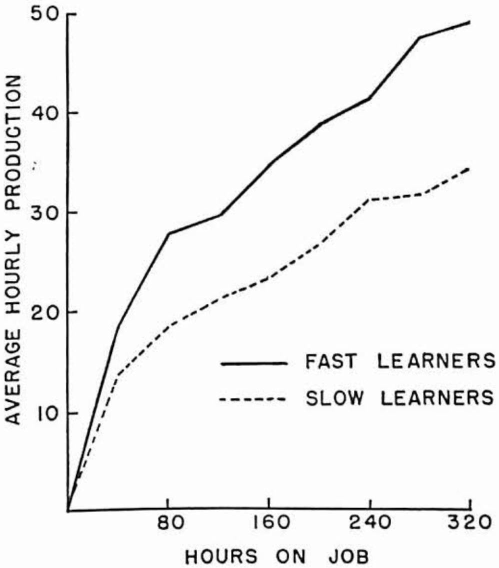


FIG. 9.11. Average hourly production in relation to hours of experience of "fast" and "slow" learners on the job of preparing spools for a loom. (Adapted from McGehee, op. cit.)

have not been graphed in terms of Bedaux units, the general principles involved in their construction are the same as those just discussed.

Data bearing on this subject reported by McGehee[27] are plotted in Fig. 9.11, which shows the average production efficiency of fast and slow learners on the job of preparing spools for a loom as a function of time on the job. The nine "fast" learners reached standard production in less than the 10.8 months average of the total group. The twelve "slow" learners required more than 10.8 months to reach standard production. It will be noted in Fig. 9.11 that the fast learners produced more than the slow learners at the end of each forty-hour period from the first throughout the eighth week. McGehee interprets these results with the statement that the supervisor who wishes "to retain only operators who would make average production in less than average learning time, could have dismissed those who composed the slow group at the end of the fifth to sixth week of training. He would still retain the majority of workers who learned the job on an average of about two months less time than the average operators in the total group of twenty-one employees." It would seem that a further use for data of the type plotted in Fig. 9.11 would be to instigate a search for differences between the fast and slow learners in methods of performing the job. Such differences, if they exist, might furnish the basis of a new training procedure that would improve most of the slow learners on this job.

---

[27] W. McGehee, "Cutting Training Waste," *Personnel Psychology*, 1948, **1**, 331–340.

Production Time Required to Do the Job. It is sometimes more convenient to measure the effectiveness of training in terms of the reduction in time required for the job as the training progresses.

An excellent example of the application of this method to the evaluation of a job-training program has been published by Greenly.[28] Fig. 9.12—reproduced from Greenly's article—shows the

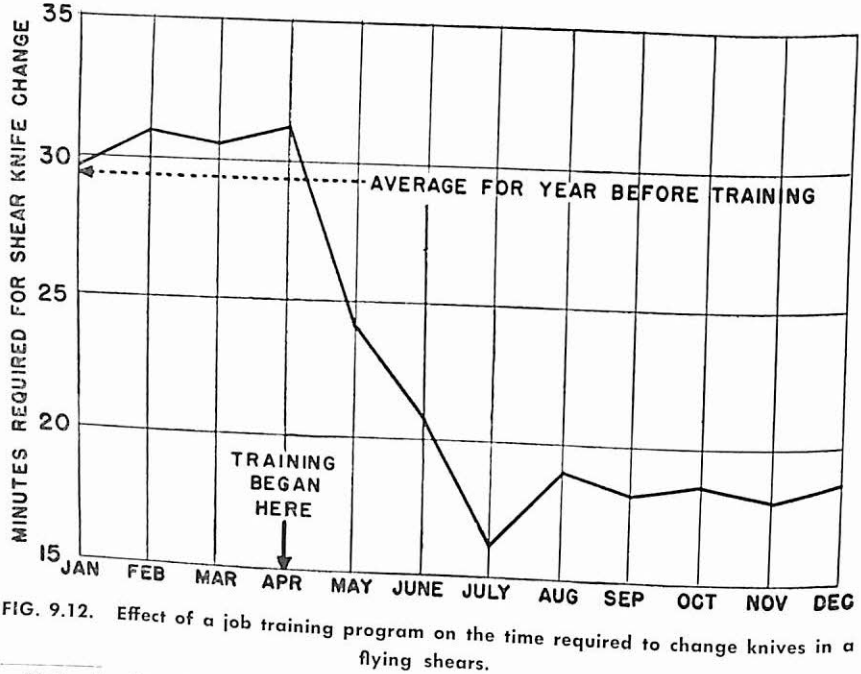


FIG. 9.12. Effect of a job training program on the time required to change knives in a flying shears.

---

[28] R. J. Greenly, "Job Training," *National Association of Manufacturers Labor Relations Bulletin, No. 35* (1941), 5–8.

reduction in time required to change the knives on a flying shears following a program of job training dealing with this operation. The graph shows that the average time required for the operation for the year preceding the installation of the training program was twenty-nine minutes. Immediately after the training was undertaken a drop in the time required for the operation is apparent. This drop continued for three successive months, at the end of which time the operation required an average time of only eighteen minutes. From a graph of this type it is a simple matter to compute the dollars-and-cents savings effected by such a job-training program. In this instance the eleven minutes saved, times the labor cost, times the average number of changes resulted in a saving of $1,740 per month or $20,880 per year.

Number of Operators Meeting Production Standard. Where production standards have been set, a measure of the effectiveness of training is the number (or percentage) of trainees who meet the standard after a given period of training. If a new training program is initiated, its value may be checked by a comparison of the percentage of operators who meet production standards after a given period of time with the corresponding percentage of operators who had met the standard after the same period of time under the previous program. This method of measuring the effectiveness of training is illustrated in Fig. 9.13, published by Lawshe.[29] In this figure the bar graphs indicate the percentages of operators meeting a given production quota after four and after eight weeks of training. The bars show the results for a group of operators in a new training program and for a group not in the program. It is apparent that operators in the program excel those who are not, both at the end of four weeks and at the end of eight weeks.

Increase in Learning Rate. Comparison of the learning curves for an operation under two sets of learning conditions provides another means of evaluating the relative effectiveness of two training programs. Such a comparison is shown in Fig. 9.14 from Lawshe.[30] In Fig. 9.14 are plotted the earnings of apprentices at the end of eight successive six-months periods under a piece work method of payment,

---

[29] C. H. Lawshe, Jr., "Eight Ways to Check the Value of a Training Program," *Factory Management and Maintenance,* 1945, **103,** 117–120.

[30] *Ibid.*

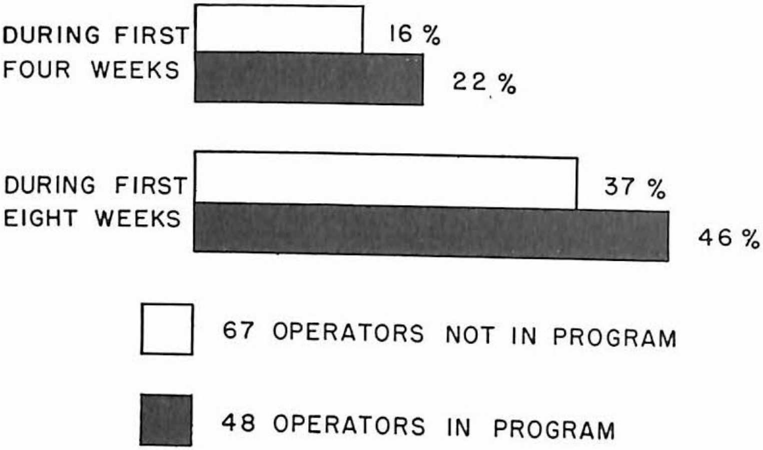


FIG. 9.13. Effect of an "induction" program upon per cent of employees engaged in a record press operation who met production standards after four and eight weeks, respectively. (Figs. 9.13 to 9.18, inclusive, courtesy Factory Management and Maintenance, from C. H. Lawshe, Jr., "Eight Ways to Check the Value of a Training Program," Factory Management and Maintenance, 1945, 103, 117–120.)

in comparison with earnings during corresponding periods of a group of apprentices paid the customary contract rates. It will be noted that under the piecework plan the apprentices are earning the same amount at the end of thirty months (five six-month periods)

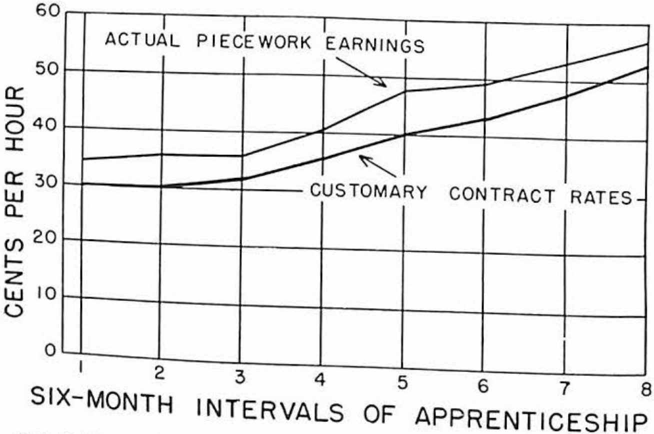


FIG. 9.14. Production of apprentices paid on an incentive plan in comparison with those paid customary contract rates.

as the apprentices under customary contract rates earn after forty-two months (seven six-month periods).

Decrease in Breakage or Use of Consumable Supplies. If two training methods result in a significant difference in use of tools, parts, and other consumable supplies, the method involving the smaller use of supplies is ordinarily the better one. This quantitative method of measuring the effectiveness of training is illustrated in Fig. 9.15.[31] The operators whose work is charted were on the job of cutting small tungsten discs on a machine using a rotary abrasive cutting wheel. One of the major costs of the operation was the use and replacement of these wheels. In Fig. 9.15 the horizontal dotted lines show the percentage of wheels broken by previous employees, who were not in the training program, at the end of eight weeks, twenty weeks, and thirty-six weeks, respectively. It is apparent from the chart that in each case the trainees broke fewer abrasive wheels than did previous operators at the end of equal periods on the job.

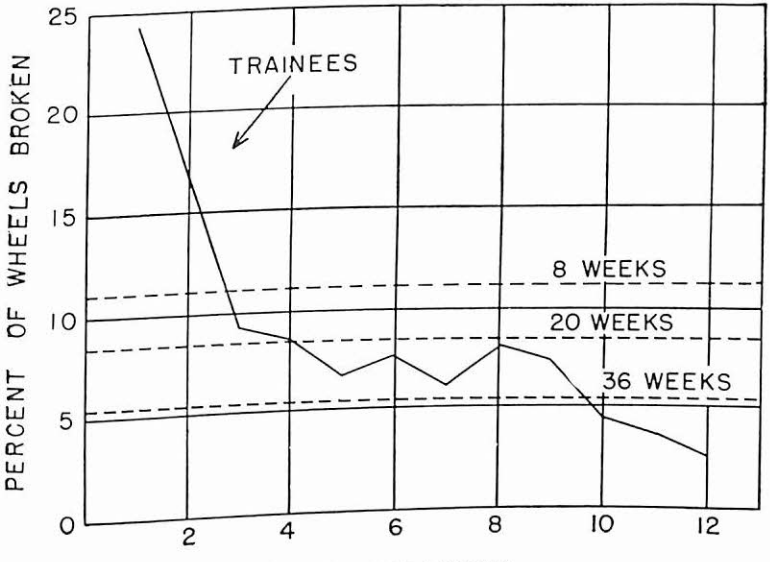


FIG. 9.15. Decrease in abrasive wheel usage of trainee group in comparison with present employees of varying amounts of experience.

---

[31] Lindahl, *op. cit.*

Reduction in Number of Accidents. In many plants and on many operations, accident prevention is one of the basic matters covered by a training program. In such situations a means of measuring the effectiveness of training may be found in a comparison of accident-experience data before and after the beginning of a new training program. This type of measurement is illustrated in Fig. 9.16 from Lawshe.[32] In this chart are plotted the percentage of employees having one or more accidents during their first thirty days on the job. The top bar is based on the record of the last seventy employees broken in "on the floor," while the lower bar is based on that of the first seventy employees to go through a new vestibule training school.

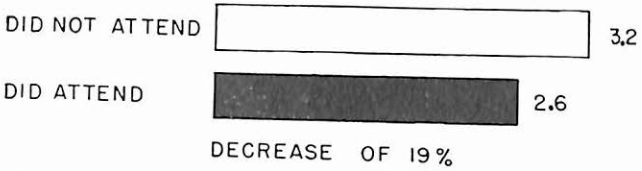


FIG. 9.16. Comparison of accident experience during first 30 days of employment of employees attending a vestibule training program and employees broken in "on the floor."

Reduction in Absenteeism. A similar measurement of the effectiveness of training may often be obtained from records of absenteeism. This method is illustrated in Fig. 9.17, also from Lawshe.[33] During each group's first thirty days of work there was an average of 5.1 days absent among the seventy operators who were trained on the floor,

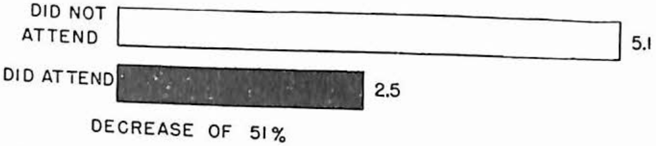


FIG. 9.17. Comparison of absenteeism during first 30 days of employment of employees attending a vestibule school and employees broken in "on the floor."

---

[32] Lawshe, *op. cit.*

[33] *Ibid.*

in contrast with an average of only 2.5 days absent among seventy operators who had vestibule training.

Reduction of Labor Turnover. The expense of training new men is usually a significant part of production costs. Other things being equal, a training program that keeps labor turnover at a minimum should be sought. A further measurement of the effectiveness of the vestibule school referred to above was obtained from data on turnover before and after the school was inaugurated. These results are graphed in Fig. 9.18, from Lawshe.[34] Other investigations, such as one reported by Kelly and Ware,[35] often show that employees who are given systematic orientation show a lower turnover rate than do employees not so inducted into their jobs.

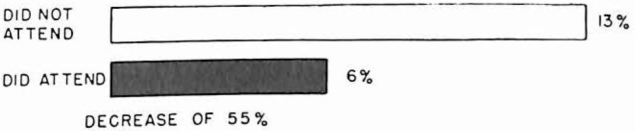


FIG. 9.18. Comparison of terminations before 30 days of employment of employees attending vestibule school and employees broken in "on the floor."

Measurement of Information. To measure the effect of training in terms of immediate monetary savings, using one or more of the indices described above, is not always possible. However, measurement may often be made of the effectiveness of training in such matters as company policy, knowledge of job, technical information, and so on, by means of specially constructed tests. For example, many industries devote a considerable amount of attention to supervisory training conferences. Such conferences frequently emphasize the importance of company policy and devote a considerable amount of time to instructing supervisors in ways of handling numerous kinds of problems, the solutions of which require a knowledge of company policy. It is not necessary to *assume* that such conferences implant the necessary or desired information in the mind of the supervisor, nor is it necessary to rely upon the supervisor's answers to such questions as, "Have these conferences improved your ability to handle situations?" A more satisfactory method of evaluating such conferences is to resort to the general principle of measurement. Such measurement can readily be made by means of specially constructed true-false or multiple-choice questions that cover thoroughly the subject-matter of the conferences and that determine in quantitative terms just what the supervisors know and do not know about the subject-matter before and after the conferences have been held. Typical examples of the kind of questions that may be used for this purpose have been published by Mapel.[36]

---

[34] *Ibid.*

[35] R. W. Kelly and H. F. Ware, "An Experiment in Group Dynamics," *Advanced Management,* 1947, **12,** 116–119.

If an employee is called out for work at 8:00 P.M. because of a vacancy in a position regularly filled by another employee, and he works until 12:00 P.M., he should be paid for:

1. 4 hours work.
2. 5 hours work.
3. 6 hours work.
4. 8 hours work.

In department "A" there are three shears: John Jones is first helper on No. 1 shear, who, due to sickness of the shearman, has acquired two weeks experience as a shearman. Several months later a vacancy exists in the shearman's position. First chance at the vacancy should be given to:

1. John Jones.
2. The first helper with the longest departmental service.
3. First helper with the greatest plant service.
4. The first helper with the longest service where combined factors are relatively equal.

Questions such as these result in a *measurement*—not an estimate or a guess—that reveals whether the men who have been trained actually have achieved an increased understanding of company policy and the union contract.

Questions of this general form may be used in still other areas of training to determine the effectiveness of instruction. Many industries have in-service training that covers technical information or content peculiar to the individual industries. The trade of pipe fitting, for example, though many of its details are common in various industries, often involves certain particular information of a technical nature that is required of the pipe fitters in a given industrial plant. If training covering this specific technical information has been given, it is only reasonable for management to ask whether the training has been effective. This can be determined by such questions as the following, taken from a typical test used in a large steel mill:

---

[36] E. B. Mapel, "Stimulating Employee Self-Improvement," *Personnel Journal*, 1941, **19**, 316–324.

> The correct pressure to use when unloading an acid tank car is
> (1) 27 (2) 36 (3) 64 (4) 90 pounds ( )
>
> The correct pressure to use when unloading an oil tank car is
> (1) 27 (2) 36 (3) 64 (4) 90 pounds ( )
>
> Oil and grease must be kept out of a pipe used for explosive materials because (1) contaminated gases do not burn so well, (2) an explosion will result, (3) it will cause clogging of the burners, (4) it makes repair work more difficult.
>
> City water lines should not be connected to mill water lines because (1) this would contaminate the city water, (2) city water is more expensive and should be used for drinking only, (3) the chlorine in the city water would damage machinery and steel product, (4) mill water is under high pressure and this would burst the city water lines.

The foregoing illustration will suffice to show that whatever may be the subject-matter of industrial training it is usually possible to devise some sort of quantitative method to measure its effect. The measurement may involve any one or several of the methods discussed, and in some instances it is, indeed, desirable to utilize several to measure the effectiveness of training. In many situations, then, management need no longer accept a training program on faith or because of someone's flowery statements about the presumed value of the program. The application of psychological measurement may make it possible to evaluate training objectively. Without such measurement, there is no way of knowing whether the training is worth what it costs, or, indeed, whether it is worth anything at all.

TYPES OF TRAINING

The various specific types of employee training programs generally can be grouped into the following areas:

1. *Orientation training*. Such training is directed toward orienting new employees (and even present employees) to the company. In particular it is directed toward providing information about the company, its products or services, its policies, organization, and so forth, and toward developing favorable attitudes on the part of employees.

2. *Vestibule training.* This is for the development of job skills, but is conducted in a separate training facility rather than at the actual job location. In vestibule training the job situation is simulated, but the entire emphasis is on training, not on production.

3. *On-the-job training.* This consists of training while the employee is actually on the job.

4. *Apprentice training.* Apprentice training typically consists of specified amounts of classroom instruction, and of shop practice under the direction of a craftsman. It is directed toward qualifying apprentices to become fully qualified craftsmen.

5. *Technical training.* This type of training is set up to provide instruction for employees in various technical areas related to their jobs, such as drafting, laboratory procedures, and so on.

6. *Supervisory and management training.* This is directed toward improving the effectiveness of supervisory and management personnel.

7. *Other training programs.* Among other training programs is the type that is used by some companies in training college graduates who have been employed upon graduation; such training frequently consists of a period of months, or even years, of experience in many of the departments of the company. Although such employees usually are assigned to jobs in whatever department they are with at the time, this variety of experience usually is thought of as training for future management responsibilities, by giving them an intimate familiarity with many phases of the company. Another type of training carried out by some companies provides opportunities for employees to take general subjects (that may, or may not, be related to their jobs), and cultural, and even recreational subjects.

In the administration of training, most training is given by the company itself, although "experts" may be brought in on certain aspects of the program. In some cases, however, a company may make arrangements with a high school, college, university, or other educational institution, to provide training for its employees. This is particularly the case with the classroom instruction of apprentice training, with technical training, and with some of the miscellaneous types of programs.

Of the several types of programs listed, the most common would seem to be on-the-job training, apprentice training, and supervisory and management training. We will therefore discuss these methods in greater detail.

On-the-job Training. In one sense, the term "on-the-job training" is a catch-all term, since it has been used in many different situations, including some in which, for practical purposes, employees have not received anything that could be called "training." Where honest-to-goodness on-the-job training is given, there is a systematic and organized effort made to train new employees, and to retrain present employees, in the job skills they are supposed to have, this being done while they are actually performing their jobs. This may be done by the supervisor, or it may be the function of a job trainer whose sole function at the time is that of training.

The sources of information for the job trainer (whether he is also a supervisor or not) are varied. In some instances it may be desirable for him to spend a period of time in the factory in which the machine itself is manufactured. By so doing he often learns many things about the operation of the machine that could be learned only with great difficulty, if at all, without this factory experience. It is often desirable for the job trainer to work in close co-operation with the industrial engineering department. By motion-study analysis it is often possible to change the work pattern so that the actual number of motions or complexity of motions in performing the job is considerably reduced. In nearly every instance it is advisable to have the trainer work as a *trainer* while he is on this job and not partly as a trainer and partly as a producer or supervisor. The illustration that will be cited in Chapter 13 on the training of inspectors typifies the difficulties often experienced when training is offered by an employee who also is responsible for certain production. Of course it is often impossible to allocate any man or men permanently to the job of training. When the training has been completed, the job trainer often returns to his former activities, either as a producer or a supervisor. But usually it is advisable for him to function as a trainer when he is doing training and as an operator or supervisor when he is not doing training, and not attempt to do both jobs at the same time.

A great deal has been said in the early part of this book concerning the magnitude and significance of individual differences in employee productivity. It was emphasized that these differences are often due to individual differences in employee capacity for the job, which is measurable by appropriate tests at the time of hiring or placement. But it should also be recognized that individual differences in employee productivity are often due in large part to differences in the training that employees have been given. Seashore[37] has pointed out that at least three basic factors may cause individual differences in human abilities. These factors are: differences in biological bases of performance, differences in the amount of previous training, and differences in the particular work methods that are adopted by each individual during the learning period. If several new employees are started simultaneously upon a new job that is somewhat complex in nature, and if these employees are not given identical and standardized instruction in the operations of that job, it is quite unlikely that all will be equally skillful in adopting the most satisfactory work methods. It is exactly as if a number of persons start simultanteously, without instruction, to learn the operation of a typewriter. Some, if not most, persons in this situation will adopt the so-called "hunt-and-peck" system, because this system is easily understood even without instruction and because it results almost at once in some evidence of accomplishment. A few, perhaps, who take the trouble to study the nature of the problem and thus become familiar with the fact that the touch system exists, and who are willing to sacrifice immediate evidence of accomplishment in favor of a learning procedure that eventually will result in considerably greater accomplishment, will not attempt the "hunt-and-peck" system at all but will start immediately to learn the touch system. After a period of time, say three months, those who learn the touch system will be many times more skillful at operating a typewriter in terms of both speed and accuracy than those who followed the more immediately apparent method of "hunting and pecking." With full recognition of the greater efficiency of the touch system, every commercial school supervises the training of students in typing so that all are taught the more desirable method. To facilitate the instruction, use is made of such devices as the blank keyboard to insure adherence to the preferred method and to avoid any relapse, even temporarily, to an undesirable method.

In many ways, performance on an industrial job is similar to typing. If employees are placed upon a new job without formal instruction in the operation of the job, they may adopt inferior methods of performing the job. They may use work methods, to use Seashore's terminology, that will become fixed in the form of work habits that result in generally low production and that later must be unlearned before proper methods can be instituted.

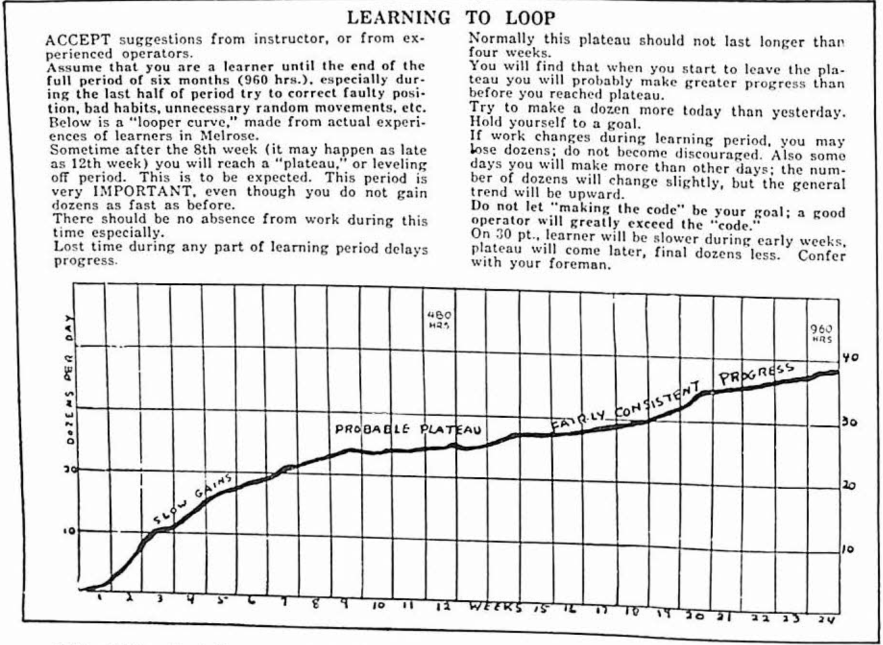
---

[37] R. H. Seashore, "Work Methods: An Often Neglected Factor Underlying Individual Differences," *Psychological Review*, 1939, 46, 123–141.

This situation can largely be offset by a program of job training organized so that every new employee is taught from the outset the proper way to do his job. The job trainer must be a man who is thoroughly acquainted with motion study and who either has conducted motion studies of the job or has available such motion studies from the industrial engineering department. He must be a man who is able to teach the proper method to the new employees. The ability to teach presupposes something more than the personal skill that he himself may possess. He must be both able and willing to explain the approved methods, be patient while methods are being understood and followed, and be willing to devote considerable energy to the encouragement of employees during those periods in the learning process when little or no improvement or accomplishment seems apparent to the learners themselves. For example, reference to Fig. 2.5 on p. 21 will show that the average looper in the hosiery industry reaches what is called a *plateau* in learning from the seventh to the twelfth month on the job. During this plateau period little if any improvement in production is attained and as a result many employees, feeling that they have attained the limit of their production and being unsatisfied with their earnings, are likely to quit the job. A considerable investment is required by the company to bring an employee up to this level of performance. If an employee decides to quit at this point, the investment is largely wasted. One of the functions of this job trainer, then, is to keep the employees on the job until this plateau is passed. To do this requires a knowledge of the evidence of the plateau and an ability to explain it to the employees.

One device used by the trainers in a hosiery plant as an aid in teaching new loopers is illustrated in Fig. 9.19 and Fig. 9.20. The facts given on these cards, which are distributed to the new employees, serve to emphasize a number of points that facilitate learning the looping operation.

The job trainer has many responsibilities in addition to instruction in the specific routine of job performance. Often he is the only



LEARNING TO LOOP

ACCEPT suggestions from instructor, or from experienced operators.

Assume that you are a learner until the end of the full period of six months (960 hrs.), especially during the last half of period try to correct faulty position, bad habits, unnecessary random movements, etc.

Below is a "looper curve," made from actual experiences of learners in Melrose.

Sometime after the 8th week (it may happen as late as 12th week) you will reach a "plateau," or leveling off period. This is to be expected. This period is very IMPORTANT, even though you do not gain dozens as fast as before.

There should be no absence from work during this time especially.

Lost time during any part of learning period delays progress.

Normally this plateau should not last longer than four weeks.

You will find that when you start to leave the plateau you will probably make greater progress than before you reached plateau.

Try to make a dozen more today than yesterday. Hold yourself to a goal.

If work changes during learning period, you may lose dozens; do not become discouraged. Also some days you will make more than other days; the number of dozens will change slightly, but the general trend will be upward.

Do not let "making the code" be your goal; a good operator will greatly exceed the "code."

On 30 pt., learner will be slower during early weeks, plateau will come later, final dozens less. Confer with your foreman.

DOZENS PER DAY
480 HRS
960 HRS
40
30
20
10
SLOW GAINS
PROBABLE PLATEAU
FAIRLY CONSISTENT PROGRESS
1 2 3 4 5 6 7 8 9 10 11 12 WEEKS 15 16 17 18 19 20 21 22 23 24

FIG. 9.19. Training material distributed to new loopers as a part of the training program of the Melrose Hosiery Company. (Courtesy J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N.C.)

man with whom the new employees are in daily contact. The impressions that they form of the company in the early stages of their employment, and that often remain with them as long as they remain employees, are often largely dependent upon their reactions to the job trainer. He must be able to explain company policy to the employees and to give them convincing reasons for the existence of these policies. He must be able to explain the wage-payment plan in operation and to show why this plan has been adopted by the company rather than some other plan employees might think is fairer. If he finds that certain phases of company policy are not understood generally or are not accepted by new employees, it is his job to relay this information to the department of industrial relations or to the other management men who are responsible for the adoption of the policies in question.

PRACTICAL SUGGESTIONS FROM EXPERIENCED OPERATORS

Establish friendly, congenial relations with instructor and fellow workers. Everyone wants to help you succeed. Do not try to learn too fast. Be patient. Learn a little at a time.

What To Do

1. POSITION. Try to relax. Some loopers urge that you sit erect, feet on floor with ankles crossed. Adjust chair to machine so that eyes meet points evenly.
2. HOLDING THE WORK. Hold loosely with thumb and index fingers of both hands. Some say use as many fingers as possible so as to hold lots of stitches in line that they may be pushed on with opposite fingers. The thumb is most important.
3. LEARN THE STITCH FIRST. Learn to put on stitches while machine is operating very slowly.
4. TURNING CORNER. Be sure to turn the first corner right and the last will be okay. Put on about third of sock, then turn corner by slipping stitches off with thumb and finger until directly in corner, then put top stitch over bottom stitch. When corner has been turned, continue to put on first side, then finish back side until complete sock is looped.
5. KEEP YOUR MIND on your work.
6. SEE HOW FAST you can pick up a sock after one sock is finished.
7. AFTER LEARNING, time yourself occasionally by dozens.
8. FORM THE HABIT, from the beginning, of inspecting your own work; at least two socks out of each dozen looped.

Some Don'ts

Don't force work on, let it slip on "naturally."

Don't pull up or down, just be sure work is on the looping machine points "good."

Don't hold work tight; this tires hands.

Don't worry about how many dozens you should make, just keep at it.

Some say don't change hands when you turn corners.

Don't put the right hand underneath the looper; this is a waste of time. (Applies to right-hand loopers.)

Don't try to hurry; make speed by working steadily. Try to gain some each day.

Gossip, bad feelings, emotional upsets interfere with production. Insofar as possible everyone at Melrose wants good, congenial work conditions in the shop.

FIG. 9.20. Further training material distributed to new loopers as a part of the training program of the Melrose Hosiery Company. (Courtesy J. D. Boyd, Personnel Director, Melrose Hosiery, High Point, N.C.)

Apprentice Training. Another area of industrial training in which still greater foresight must ordinarily be exercised is the training of skilled tradesmen. This training is usually accomplished by means of an apprentice program. Except for college-trained technical men there is perhaps no other single group of industrial employees who require so extensive a period of training for expert performance as do skilled tradesmen. The federal committee on apprenticeship training specifies that the training period for apprentices shall be not less than 4,000 hours, and that the shop instruction shall be accompanied by not less than 144 hours of instruction under public authorities.[38] These stipulations reflect the well-substantiated judgment of those experienced in this field that adequate training of skilled tradesmen requires an intensive and extensive instructional period. These facts make it more necessary for industry to look into the future in the training of skilled tradesmen than in the training of many other types of industrial employees. The various details in the administration of an apprentice program, together with the contractual agreement between the apprentice and the employer, are not of primary concern to the technical psychologist, and therefore will not be discussed here. Psychology, however, may be definitely applied in the use of aptitude and qualification tests for the selection of boys to receive apprentice training. While mass production in industry has reduced the proportion of employees who must be skilled tradesmen, the design, construction, and maintenance of machines has increased so much in complexity that a very high level of capacity is required of the boy who is to become a skilled tradesman. Fig. 2.1 in Chapter 2 shows that the range of ability among applicants for apprenticeships is large indeed. Any program of apprentice training that does not provide an adequate means of selecting trainees according to their capacity to benefit from the training is likely to result not only in a failure to develop expert tradesmen but also in the waste of a considerable company investment in the training given. Those responsible for the administration of an apprentice program should therefore work in close collaboration with the testing division of the personnel office so that trainees will be selected who are able to take full advantage of the opportunity offered.

Supervisory and Management Training. Probably the most important new emphasis in training in recent years is in the area of supervisory and management training. Although some companies have both supervisory training and management training programs, the line between them is a thin one. The differences are essentially those of degree rather than of kind.

As pointed out by Fryer, Feinberg, and Zalkind,[39] the basic training for both managers and supervisors deals with principles of scientific management relative to organization, planning, direction, and control.

---

[38] *The National Apprenticeship Program* (War Manpower Commission Bureau of Training, Apprentice-Training Service, 1943).

[39] Fryer, Feinberg, and Zalkind, *op. cit.*, p. 174.

The increased emphasis on supervisory and management training in recent years has paralleled a shift in the duties and responsibilities of supervisory and management personnel. The job of foreman, in particular, has undergone a change. In years gone by, a foreman typically had responsibility for employment and discharge of employees, setting wage rates, determining methods of work, quality control, training, and discipline (often administered arbitrarily, and with a heavy hand). Many of these functions have, in part or entirely, been taken over by staff personnel. In the meantime, additional responsibilities have fallen to the lot of the foreman, such as interpreting company policies and regulations, social legislation, union contracts, and the like; handling grievances and other matters relating to the union if there is one; serving as a communications link between management and employees; and having a key role in the development and maintenance of good human relations with the employees. There have been related changes in the functions of other levels of management, as well. Present-day supervisory and management training programs are directed toward the development of such personnel to carry out their responsibilities more effectively.

Nature of supervisory and management training. A schematic diagram representing the major aims and objectives of supervisory training, and the typical content and methods of such training, is shown in Fig. 9.21. The objectives of orientation, management skills, human relations skills, and technical skills are achieved by training of different types, as indicated by the content of the training shown for these distinct objectives. For any given organization, the *training needs* will of course be somewhat unique, so the training program should be directed toward fulfilling those specific needs. Fig. 9.21 is simply a generalized chart representing various aspects of supervisory training.

The most frequently used methods of supervisory and management training include conferences, discussions, case studies, and role playing. Role playing involves dramatizing work situations by having those in the group play the "roles" of individuals in the situation being dramatized. For example, two individuals might play the "roles" of a supervisor and a subordinate in a situation in which the subordinate is to be reprimanded by the supervisor. Case studies, in

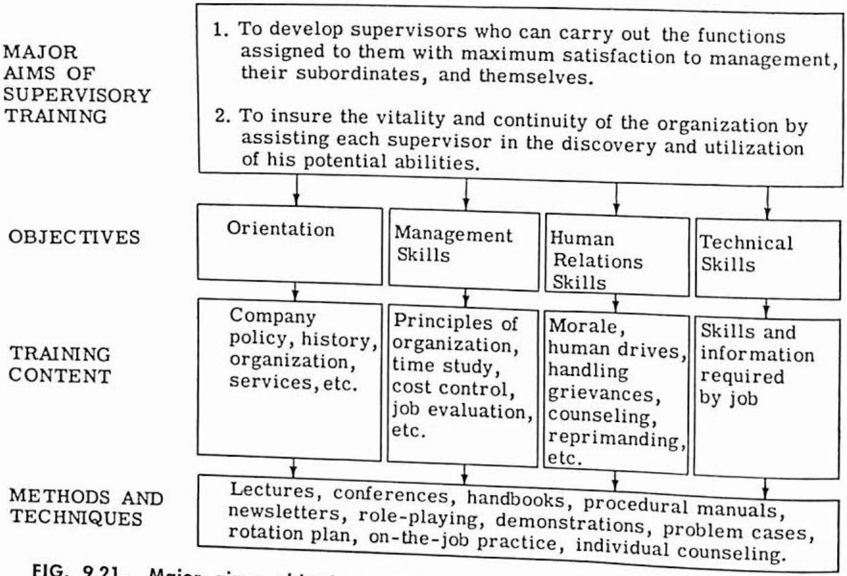

MAJOR AIMS OF SUPERVISORY TRAINING
1. To develop supervisors who can carry out the functions assigned to them with maximum satisfaction to management, their subordinates, and themselves.
2. To insure the vitality and continuity of the organization by assisting each supervisor in the discovery and utilization of his potential abilities.
OBJECTIVES
Orientation
Management Skills
Human Relations Skills
Technical Skills
TRAINING CONTENT
Company policy, history, organization, services, etc.
Principles of organization, time study, cost control, job evaluation, etc.
Morale, human drives, handling grievances, counseling, reprimanding, etc.
Skills and information required by job
METHODS AND TECHNIQUES
Lectures, conferences, handbooks, procedural manuals, newsletters, role-playing, demonstrations, problem cases, rotation plan, on-the-job practice, individual counseling.

FIG. 9.21. Major aims, objectives, content, and methods of supervisory training. (Adapted from Earl G. Planty, William S. McCord and Carlos A. Efferson, Training Employees and Managers, The Ronald Press, 1948.)

turn, are problem situations (frequently based on actual personnel problems such as those that have occurred in the relations between a supervisor and a subordinate) that are described to the training group. The training group then discusses the case, and works out a "solution."

The selection of a method of training should be appropriate to the purpose, of course. For example, a supervisory training conference should not be expected to achieve success in covering topics outside the realm of experience of the group attending the conference. When new subject-matter is to be covered, the information meeting is much more appropriate. As Allen[40] has summarized the matter, conferences as such are of value only if the supervisors all "have a fairly common level of information and knowledge about the situation." If not, the information meeting is more appropriate.

Evaluating effectiveness of supervisory training. It has already been emphasized that some means of measuring the effectiveness of training should be employed. A typical method of measuring the effectiveness of training in company policy is illustrated by supervisory case problems, with related true-false questions, such as the one reproduced in Table 9.6 on pp. 298–299.

---

[40] L. A. Allen, "Inform or Confer?" *Personnel*, 1950, **26**, 448–450.

The Joe Blake case (the name is fictitious) is an actual case that arose in one plant during one of the conferences on company policy. The answers to the 20 true-false questions obtained from a group of supervisors not only furnish management with a means of knowing how well supervisors are prepared to handle this type of problem but also indicate what areas of company policy should be covered in the supervisory training conferences in order to achieve a unified understanding of company policy on the part of supervision. A series of 20 cases, with 20 true-false questions on each case—from which Case I was taken—furnishes the management with a 400-item true-false test with which to measure how much the supervisors know about the administration of company policy. True-false or objective questionnaires, formerly considered purely academic in application, have been found to be serviceable as measuring instruments in practical industrial situations.

Similar use may be made of multiple-choice or completion form questions dealing with company policy or the contract with the union. An example of the latter has been published by Pond and O'Keefe.[41] Sample questions from their questionnaire are:

1. Except in emergencies, at least . . . . . . days' notice shall be given by the Company when changes in hours or assignments to shifts are deemed necessary by the Company.
2. An employee called back to work for a period not continuous with his regular tour of duty shall be paid at overtime rates for not less than . . . . . . hours.
3. Except in emergencies, an employee who is required to work on his holiday shall be notified . . . . . . days in advance of such holiday.
4. Seniority rights of employees laid off shall terminate if they are not rehired within . . . . . . months after layoff.
5. When the Company decides to transfer employees from one shop to another, such bids shall be posted for . . . . . . days.

In the above examples, the "criteria" of training effectiveness consisted of tests or questionnaires given to the supervisors at the completion of training. Such criteria, however, are not "ultimate"

---

[41] S. A. Pond and R. W. O'Keefe, "The Contract Questionnaire," *Personnel,* 1950, **26,** 304–306.

TABLE 9.6

A Test Used in Measuring What Supervisors Know About Company Policy

CASE NUMBER 1
SUPERVISORY PROBLEMS

Joe Blake is a tractor operator in No. 8 shipping department. He has worked for the company six months and had a clean record until a week ago. Before coming to this job he worked for a neighboring company and was highly recommended by them. A week ago his tractor became overheated because of a dry radiator and had to be sent to the shop for repairs. Both shipping and repair departments were behind schedule and the resulting delay due to the absence of this tractor caused considerable confusion and loss.

One of the tractor rules states that each operator shall inspect his machine thoroughly before work is begun, and also before turning the machine over to the next operator. Joe Blake, like all other operators in the department, does not make the first inspection but, according to agreement among the tractor operators, makes a careful inspection at the end of the turn so the following man finds the machine ready to go.

After investigating, the foreman reported the incident as an act of carelessness and a clear violation of the plant rules, and consequently penalized Joe by giving him a one-week lay-off. Joe reminded the foreman that it was a well-known and tolerated practice to omit the first inspection and brought in the previous operator who insisted that he had filled the radiator just before the breakdown. Being a new man, Joe did not feel it wise to argue further and accepted the penalty with good grace.

During the following week the rumor arose and was substantiated that the tractor breakdown was caused by a group of six packagers, who as a joke, drained the radiator of Joe's tractor while he was away in the lavatory. These details and the names of the men involved were known to nearly everyone in the department.

CASE NUMBER 1

Plant. . . . .
Dept. . . . .

INDUSTRIAL RELATIONS—CARNEGIE-ILLINOIS STEEL CORPORATION

The following statements apply to this situation. Study each one carefully. If you agree, put a circle around the "T" following the statement. If you do NOT agree, encircle the "F." If you are in doubt, encircle the double "??." Be sure to show your response clearly.

1. The six packagers should be given penalties equal to Joe Blake's. T ?? F
2. The foreman should apologize to Joe and void the rule violation report. T ?? F
3. The foreman should let the report stand as it is, since responsibility for the continuous operation of the tractor is Joe's, regardless of the action of others. T ?? F
4. Joe should be given enough overtime work to make up for his week's loss. T ?? F

5. The foreman should be fired for allowing the inspection rule to be violated openly. T ?? F

6. The foreman should not change his action unless Joe takes the matter up with his grievance committee and threatens to demand redress. T ?? F

7. The departmental supervisor should take advantage of the situation by holding a meeting with the men, pointing out the waste and loss caused by a "practical joke" and a violation of plant rules. T ?? F

8. Joe should be transferred to another department and other steps should be taken to "hush up" the foreman's action. T ?? F

9. The company should refuse to recognize the act of the packagers since delayed recognition of that aspect of the case might be held against the company at a latter date as an example of management's arbitrary policies carried out without proper investigation. T ?? F

10. A general survey should be made to see that plant rules are being obeyed to the letter throughout the department. T ?? F

11. All the tractor operators in the department should be penalized for wilful neglect and carelessness because of their practice in regard to the inspection rule. T ?? F

12. The foreman should protect his men by explaining to his supervisor that he had known of and tolerated the limited inspection practice. T ?? F

13. The foreman was not responsible for the limited inspection practice even though he knew of the rule violation. T ?? F

14. The foreman should conceal the matter from his supervisor, if possible, since admission of his tolerance of a rule violation would serve only to increase friction and undermine his authority. T ?? F

15. The company should not permit the foreman and superintendent to admit that an error of hasty misjudgment was made, since that would lower employee morale. T ?? F

16. The tractor operators as a group had no responsibility in this case because the inspection practice was tacitly approved by the foreman who is management's representative. T ?? F

17. The responsibility for the occurrence rests solely with the six packagers. T ?? F

18. This is one of those situations in which the responsibility cannot be fixed. T ?? F

19. Blame for the occurrence rests solely with the foreman. T ?? F

20. A foreman should be certain of his information and of all the facts before he takes action against an employee. T ?? F

Source: R. J. Greenly, "The Case Method in Supervisory Training," American Management Association Personnel Series No. 56 (1942), 30–35.

criteria. The ultimate criterion of the effectiveness of a supervisory or management training program would consist of some measure of the effectiveness of the *group* of persons under the supervision or management of those receiving training. Such a measure might be in terms of group productivity or group morale. If supervisory or management training can be demonstrated to result in group effectiveness (presumably because of improved supervision or management), the training certainly can be said to have paid off. Since such criteria are hard to obtain, however, it is usually the practice to use more readily available criteria, such as the test mentioned above.

Carry-over of training to work situation. To be effective, a supervisory or management training program should result in some carry-over improvement in the effectiveness of the group supervised or managed. There is some evidence, however, that the extent to which such training does "rub off" on the group is in large part determined by the total leadership "climate" of the organization. This implication comes from the results of a study by Fleishman[42] in which he attempted to evaluate a leadership training course for 122 foremen after the foremen returned to the industrial situation.

The basic data for the study came from the series of questionnaires listed below:

Completed by *foremen*:

1. Foreman's Leadership Opinion Questionnaire.
2. Supervisory Behavior Description.
3. What Your Boss Expects of You.

Completed by *work groups* supervised by foremen:

1. Foreman Behavior Description.
2. How You Expect an Ideal Foreman to Act.

Completed by *bosses* of foremen:

1. Leadership Opinion Questionnaire.
2. What You Expect of Your Foreman.

These questionnaires were scored for the following factors:

1. "Consideration" (extent to which the leader was considerate of the feelings of those under him—the "human relations" aspects of group leadership).

---

[42] E. A. Fleishman, "Leadership Climate, Human Relations Training, and Supervisory Behavior," *Personnel Psychology*, 1953, **6**, 205–222.

2. "Initiating Structure" (the extent to which the supervisor plans, schedules, criticizes, initiates ideas, organizes the work, and otherwise systematizes or structures the work to be done).

A follow-up of the trained foremen was carried out over a period of 39 months. Without describing the results or conclusions in detail, one important finding will be reported. It was found that the kind of *supervisor* ("leadership climate") under whom the foreman operated had an influence on the *foreman's* attitudes and behavior. Those foremen who operated under a supervisor who was "considerate" tended also to be more considerate of their workers than did those foremen who operated under supervisors who were less "considerate." In turn, those foremen who worked under supervisors who "structured" the work situation (stressed deadlines, assigned people to specific tasks, and so forth) also tended to perform in the same way. This is something like a chain reaction; the attitudes and methods of operation of a supervisor are transmitted also to the foremen whom he supervises.

One implication of this in connection with supervisory and management training is that the training of first-line or intermediate-level supervisors and managers probably will have little permanent effect on their attitudes or behavior if they subsequently work under the direction of supervisors who create an undesirable "leadership climate." The corollary of this is that management training should begin at the top level of management. If, through training of top-level management personnel, their attitudes and behavior can be so modified that they create a favorable "leadership climate," it is then reasonable to expect that corresponding training of intermediate and lower levels of management will more likely result in desirable changes in their attitudes and behavior.

10

THE MEASUREMENT OF ATTITUDES AND MORALE

Management has a growing interest in employee attitudes, opinions, and morale. This interest is reflected in the fact that many companies conduct attitude and opinion surveys, provide "human relations" training for supervisors, sponsor recreation programs for employees, publish employee newspapers, and take other actions that are intended to create favorable attitudes on the part of employees.

This concern on the part of management probably can be attributed in part to a general trend toward greater recognition of social responsibilities by industry. In part it can probably be attributed to the belief on the part of management that employees with favorable attitudes toward their company generally are "better" employees in some respect, such as being more productive, or having lower turnover rates. There is some evidence to support this belief, but a discussion of this will be deferred until later.

Before considering some of the factors that affect attitudes and morale let us discuss their measurement. Even before doing this, however, we should crystallize some of the concepts that we will be discussing.

THE MEANING OF TERMS

Attitudes. The term *attitude* has various connotations; we will use it in the sense defined by Maier.[1] He defines an *attitude* as a kind of mental set. It represents a predisposition to form certain opinions. To put it another way, it is a frame of reference that influences the individual's views or opinions on various topics, and that influences his behavior.

---

[1] N. R. F. Maier, *Psychology in Industry* (2nd ed.; Houghton Mifflin Co., 1955).

Attitudes are formed through experience, which means that they are learned. Once a person has developed a particular attitude, it may be difficult for him to determine how he acquired it. In fact, an individual may not be consciously aware of his own fundamental attitudes. Whether a person's attitudes are based on rational considerations and factual information, or whether they have a strong emotional bias, has little bearing on the *effect* of the attitudes on the person's thinking or behavior. In either case the factor that affects behavior is the *attitude, not* the consideration of whether it is or is not a rational attitude.

Although the attitudes of people tend to be relatively stable, they can be modified, at least to some degree. Since they can be changed, it is more fruitful to measure them than would be the case if they were firmly fixed, once and for all.

Opinions. An *opinion*, on the other hand, is an expression of an evaluative judgment or point of view regarding a specific topic or subject. Thus, an attitude is somewhat generalized (such as liking or not liking a person's supervisor), whereas an opinion typically is an interpretation regarding a specific matter (such as saying that the boss plays favorites in granting overtime). Opinions, however, typically are influenced by the more generalized attitude. The facts or observations within an individual's experiences are *interpreted* in the light of his *attitudes*. Thus, if a foreman calls the attention of his work group to the fact that some of the safety rules have been violated, one person (who has an "unfavorable" attitude toward the foreman) might later express the opinion to one of his buddies that the foreman is "—just picking on us"; another person (who has a "favorable" attitude toward the foreman) might later express the opinion that the foreman is simply "—trying to keep us from getting our fingers cut off."

Morale. Whereas *attitudes* and *opinions* are characteristic of *individuals*, the term *morale* generally has an implication of *group* reactions. In other words, the morale of a group depends on the interactions among individuals in the group, such as a football team, a military unit, or a work group. It is akin to the common notion of team spirit.

MEASUREMENT OF ATTITUDES AND OPINIONS

Attitudes, opinions, and morale cannot be weighed on a scale, like a pound of butter, nor measured with a rule, like a strip of carpet. Yet, though they are intangible, they can be measured. Thorndike long ago pointed out that whatever exists, exists in some amount, and whatever exists in some amount can be measured (though this is sometimes admittedly difficult). Certainly differences in employee attitudes do exist, and, further, they can have implications for the effectiveness of the individuals within an organization, and for the organization as a whole. The morale of the people within a group, for example, can make or break a business enterprise or a military campaign.

There are various ways in which it is possible for supervisors and managers to get some inkling of the attitudes and opinions of individuals and of the morale of a group. Some of these methods are very informal and unsystematic, such as interpreting chance remarks of individuals, the behavior of individuals and work groups, and the reported "feelings" of supervisors. The lack of systematic measurement with such methods, however, leaves much to be desired. It should be added, nevertheless, that such clues should not be completely discounted. A sensitive foreman, for example, frequently can get a "feel" with respect to whether the morale of his group is high or low, even though he cannot say *how* high or low.

Another method of getting some inkling about employee morale is by analyzing changes in such factors as turnover rate, absenteeism, and production level. These clues, however, are at best only indirect, and they can also be influenced by factors *other* than morale.

For many purposes it is desirable to obtain some more adequate, and direct, measure of attitudes, opinions, and morale. Various methods have been developed for doing this, some of these consisting of questionnaires which are filled out anonymously by employees.

When questionnaires are used, the employees are asked to record their responses to the questions, and then to drop the completed questionnaire in a ballot box or to mail it to a specified address. Frequently such surveys are conducted by independent organizations such as a university or a management consulting firm. Having the survey conducted by such organizations gives the employees complete assurance that their responses will not be held against them, individually.

With anonymous questionnaires, no one knows how any particular employee answered the questionnaire, but from the responses of a group, collectively, it is possible to obtain an impression of general employee reaction. There are two general types of questionnaires in common usage. One is sometimes called an *attitude scale*, and the other an *opinion survey*.

Attitude Scales. The specific questions of typical attitude scales may call for an individual to give his "opinion" about some aspect of the company, but these "opinions," collectively, are used to obtain some general attitude level for the (anonymous) individual represented by each questionnaire. The average of the attitude scores for the employees within a group (job, department, company, or whatever), in turn, is used as an index of "morale" of the group. There are various types of attitude scales, but they all serve essentially the same purposes.

A typical example of an attitude scale that may be used to determine the general attitude of employees toward their company is shown in Table 10.1. This scale is taken from an article by Bergen.[2] The theory of such scales and detailed instructions for their construction have been described by Thurstone and Chave.[3] This same

TABLE 10.1

Statements Used in Bergen's Scale for Measurement of Attitude of Employees Toward Their Company

	Scale Value
I am made to feel that I am really a part of this organization	9.72
I can feel reasonably sure of holding my job as long as I do good work	8.33
I can usually find out how I stand with my boss.	7.00
On the whole, the company treats us about as well as we deserve.	6.60
I think training in better ways of doing the job should be given to all employees of the company.	4.72
I have never understood just what the company personnel policy is	4.06
In my job I don't get any chance to use my experience.	3.18
I can never find out how I stand with my boss.	2.77
A large number of the employees would leave here if they could get as good jobs elsewhere.	1.67
I think the company's policy is to pay employees just as little as it can get away with.	.80

---

[2] H. B. Bergen, "Finding Out What Employees Are Thinking," *The Conference Board Management Record*, April 1939.

[3] L. L. Thurstone and E. J. Chave, *The Measurement of Attitude* (Chicago, University of Chicago Press, 1929).

procedure was mentioned in Chapter 8 on "Industrial Merit Rating," since it also lends itself to use in the development of weighted check list merit rating systems.

In the development of attitude scales by this method, the first step is to write out a large number of statements, perhaps a hundred or more, each of which expresses a viewpoint of some kind toward the company. An effort should be made to have these statements express all possible viewpoints from extremely favorable to extremely unfavorable. Each of these statements is typed on a separate slip of paper and a judge is asked to place each statement in one of several piles (usually 7, 9, or 11), ranging from statements judged to express the *least* favorable viewpoints (placed in pile 1), to statements judged to express the *most* favorable viewpoints (placed in the top pile, 7, 9, or 11). Statements judged to express varying degrees of favorableness in between these extremes are placed in the piles that are judged best to characterize their relative degrees of favorableness. A record is kept of the category in which each judge classifies each of the many statements.

Many judges are used in this process, sometimes as many as 100 or more. It should be emphasized that these judges are assisting in the construction of the scale. They are not having their own attitudes measured. To measure attitudes with the scale is not possible until the scale has been constructed, and the allocation of statements to the several piles is a part of the process of constructing the scale.

The purpose of the allocation is to determine the scale values of the various statements. If all judges tend to place a statement in piles toward the favorable end of the continuum, we may safely conclude that that statement expresses a favorable attitude toward the company. If a statement is generally placed by the judges in piles toward the unfavorable end of the series, we may likewise conclude that an unfavorable attitude is expressed by that particular statement. The piles are numbered from the unfavorable to the favorable end of the continuum. The number of times each statement is placed in each pile is determined, and a computation made to determine the average location of the statement by the judges. From this calculation, the scale value of the statement is determined, as shown in the example in Table 10.1. It should be added that an analysis is also made of the consistency of the judgments for each statement. Statements that are placed by all judges in one or in a limited number of categories are the statements that have the greatest degree of stability. Statements that are scattered by the judges over several categories are eliminated. By starting with many more statements than need be retained for the final scale, it is usually possible to identify 10 to 15 statements that are spread over the entire range, and that have been consistently judged by the judges. The statements so selected, together with their scale values as shown in Table 10.1, comprise the final material for the attitudé scale.

One might think that the attitudes or feelings of the judges who are used in the construction of the scale would have an effect on the scale values obtained. In other words, it might be felt that one set of scale values for a series of statements might be obtained if the judges were, in general, *favorable* toward the company, while a different set of values might be obtained if the judges in general were unfavorable or indifferent toward the company. This possibility has been subjected to experimental test, and it has been found that the attitudes of the judges do not significantly affect the scale values obtained from them.[4] This fact increases the possibilities of the use for such scales in industry because it often happens that the persons who are most conveniently available for use as judges in constructing a scale may be more favorably disposed toward the company than certain groups of employees with whom the scale is to be used after it is constructed.

In the practical administration of an attitude scale, statements are printed on a sheet in random order, without the scale values appearing in Table 10.1. Each employee is given one of these sheets and is requested to check all statements that he agrees with or believes to be true. The sheets are then turned in without being signed. The attitude of an employee toward the company usually is defined as the average or median scale value of the statements he has checked. If the average is to be used, an employee checking statements 1, 3, and 5 of those shown in Table 10.1 would have an attitude score of:

$$\frac{9.72 + 7.00 + 4.72}{3} \text{ or } 7.15$$

---

[4] E. D. Hinckley, "The Influence of Individual Opinion on the Construction of an Attitude Scale," *Journal of Social Psychology*, 1932, 3, 283–296.

On a scale of ten (ten being the most favorable end and zero the least favorable end) an attitude of 7.15 would be one somewhat toward the favorable end of the scale. On the other hand, an employee checking statements 7, 8, and 10 would have an attitude score represented by:

$$\frac{3.18 + 2.77 + .80}{3} \text{ or } 2.25$$

This would be a much less favorable attitude toward the company than the one described above.

Mention was made in Chapter 8 of a possible fallacy in the use of the median in merit rating, as pointed out by Jurgensen.[5] This same fallacy also can occur in the use of the Thurstone type of scale in attitude measurement. The solution is described in Chapter 8 (see pp. 216–218).

The attitude scale reproduced in Table 10.1 is by no means the only set of statements that might be constructed to fulfill the requirements of a suitable scale. Uhrbrock,[6] in a comprehensive study of the attitudes of 3,934 factory workers, 96 clerical workers, and 400 foremen, used a different scale, but one constructed as explained above to measure the general attitude of employees toward the company. Some typical statements, with their scale values from Uhrbrock's scale, are reproduced in Table 10.2.

The fact that scales such as those illustrated in Tables 10.1 and 10.2 may be used independently to measure the attitude of employees toward their company illustrates several points concerning the use of such scales. It will be noticed that the statements comprising the two scales are not the same, yet both scales measure essentially the same thing. Thus, the use of such scales does not require including any specific statements that, from an industrial relations viewpoint, might seem undesirable. A scale can be "tailor-made" for a given plant and deliberately kept free of statements that may seem to "pack dynamite" without sacrificing the validity of the scale. The use of different statements in scales measuring the same attitude also

---

[5] C. E. Jurgensen, "A Fallacy in the Use of Median Scale Values in Employee Check List," *Journal of Applied Psychology*, 1949, 33, 56–58.

[6] R. S. Uhrbrock, "Attitudes of 4,430 Employees," *Journal of Social Psychology*, 1934, 5, 365–377.

TABLE 10.2

Statements Used in Uhrbrock's Scale for Measuring Attitude of Employees Toward Their Company

Statement	*Scale Value*
I think this company treats its employees better than any other company does	10.4
If I had to do it over again I'd still work for this company	9.5
They don't play favorites in this company	9.3
A man can get ahead in this company if he tries	8.9
I have as much confidence in the company physician as I do in my own doctor	8.7
The company is sincere in wanting to know what its employees think about it	8.5
A wage incentive plan offers a just reward for the faster worker	7.9
On the whole the company treats us about as well as we deserve	7.4
I think a man should go to the hospital for even a scratch, as it may stop blood poisoning	6.3
I believe accidents will happen no matter what you do about them	5.4
The workers put as much over on the company as the company puts over on them	5.1
The company does too much welfare work	4.4
Soldiering on the job is increasing	4.1
I do not think applicants for employment are treated courteously	3.6
I believe many good suggestions are killed by the bosses	3.2
My boss gives all the breaks to his lodge and church friends	2.9
I think the company goes outside to fill good jobs instead of promoting men who are here	2.5
You've got to have "pull" with certain people around here to get ahead	2.1
In the long run this company will "put it over" on you	1.5
The pay in this company is terrible	1.0
An honest man fails in this company	0.8

facilitates checking results by a repeat test in order to be sure of conclusions reached and to measure the effectiveness of systematic company efforts to improve employee morale.

Likert[7] has developed a method of measuring attitudes that is somewhat simpler than the Thurstone method, and avoids the necessity of using unfavorable statements. Managements often object to the use of such statements. Likert's method moreover does not require the use of judges in scaling the statements. While a number of different procedures were tried and compared, the simplest method described by Likert was found to give results that correlated very highly with more complex methods. This method consists in printing after each statement five degrees of approval, and asking the person taking the scale to check one of the five. A typical statement from Likert's Internationalism Scale is as follows:

---

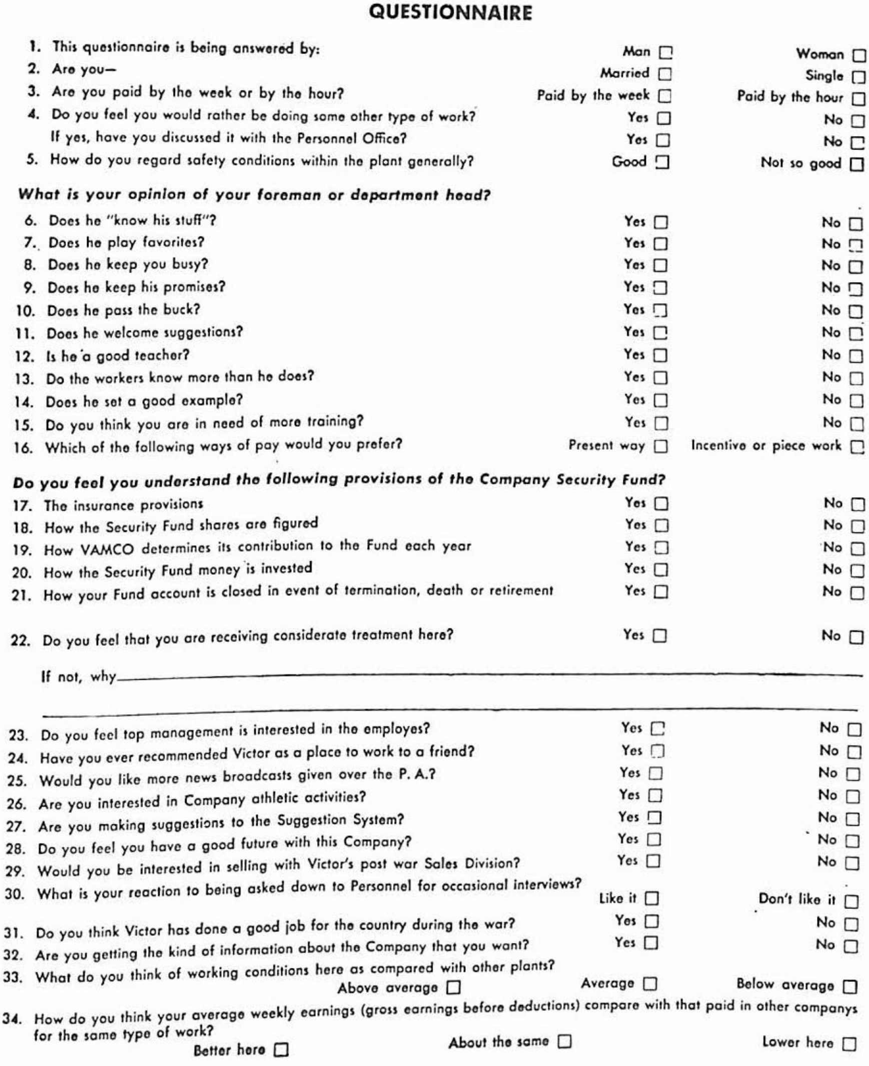
[7] R. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, 1932, No. 140, 55 pp.

The United States should have the largest military and naval air fleets in the world. Strongly approve ( ) Approve ( ) Undecided ( ) Disapprove ( ) Strongly disapprove ( )

These responses are scored from 1 (strongly approve) to 5 (strongly disapprove). By using approximately twenty-five such statements, each followed by five degrees of approval, the sum of the scores of the responses checked constitute a measure of one's attitude toward the object of the scale. Since this method can be followed by using only statements that vary between indifference (the central point on an attitude continuum) and the favorable end of the scale, the use of unfavorable statements in the scale itself is avoided. Likert has shown that this method of measuring attitude correlates approximately .80 with the more tedious Thurstone method.

Attitude scales are not perfect instruments for the registering of employee feelings, but they are considerably better than guessing or relying on chance (and often biased) individual reports.

Opinion Surveys. Attitude scales of the type described in the preceding section help to measure the attitudes of individuals; by summarizing data for all employees within a group, such scales can be used to quantify the "morale" of employee groups. Comparisons then can be made from one department to another. However, although such scales can be useful in indicating the relative level of morale of employee groups, they do not enable a management to identify specific factors that may be sources of employee unrest or dissatisfaction. This information can be obtained by means of another type of questionnaire, one that asks for opinions on specific matters, such as company policies or working conditions. This type of questionnaire should also be used without reference to any employee's name, and great care should be taken to insure the anonymous feature of the survey. An example of a questionnaire used in an employee survey of this type conducted by the Victor Adding Machine Company of Chicago is shown in Fig. 10.1. The following steps were taken in conducting the survey employing this questionnaire:



QUESTIONNAIRE

1. This questionnaire is being answered by:	Man ☐	Woman ☐
2. Are you—	Married ☐	Single ☐
3. Are you paid by the week or by the hour?	Paid by the week ☐	Paid by the hour ☐
4. Do you feel you would rather be doing some other type of work?	Yes ☐	No ☐
If yes, have you discussed it with the Personnel Office?	Yes ☐	No ☐
5. How do you regard safety conditions within the plant generally?	Good ☐	Not so good ☐
What is your opinion of your foreman or department head?		
6. Does he "know his stuff"?	Yes ☐	No ☐
7. Does he play favorites?	Yes ☐	No ☐
8. Does he keep you busy?	Yes ☐	No ☐
9. Does he keep his promises?	Yes ☐	No ☐
10. Does he pass the buck?	Yes ☐	No ☐
11. Does he welcome suggestions?	Yes ☐	No ☐
12. Is he a good teacher?	Yes ☐	No ☐
13. Do the workers know more than he does?	Yes ☐	No ☐
14. Does he set a good example?	Yes ☐	No ☐
15. Do you think you are in need of more training?	Yes ☐	No ☐
16. Which of the following ways of pay would you prefer?	Present way ☐	Incentive or piece work ☐
Do you feel you understand the following provisions of the Company Security Fund?		
17. The insurance provisions	Yes ☐	No ☐
18. How the Security Fund shares are figured	Yes ☐	No ☐
19. How VAMCO determines its contribution to the Fund each year	Yes ☐	No ☐
20. How the Security Fund money is invested	Yes ☐	No ☐
21. How your Fund account is closed in event of termination, death or retirement	Yes ☐	No ☐
22. Do you feel that you are receiving considerate treatment here?	Yes ☐	No ☐
If not, why__________		
23. Do you feel top management is interested in the employes?	Yes ☐	No ☐
24. Have you ever recommended Victor as a place to work to a friend?	Yes ☐	No ☐
25. Would you like more news broadcasts given over the P. A.?	Yes ☐	No ☐
26. Are you interested in Company athletic activities?	Yes ☐	No ☐
27. Are you making suggestions to the Suggestion System?	Yes ☐	No ☐
28. Do you feel you have a good future with this Company?	Yes ☐	No ☐
29. Would you be interested in selling with Victor's post war Sales Division?	Yes ☐	No ☐
30. What is your reaction to being asked down to Personnel for occasional interviews?	Like it ☐	Don't like it ☐
31. Do you think Victor has done a good job for the country during the war?	Yes ☐	No ☐
32. Are you getting the kind of information about the Company that you want?	Yes ☐	No ☐
33. What do you think of working conditions here as compared with other plants? Above average ☐	Average ☐	Below average ☐
34. How do you think your average weekly earnings (gross earnings before deductions) compare with that paid in other companys for the same type of work? Better here ☐	About the same ☐	Lower here ☐

FIG. 10.1. Part of a questionnaire used in an employee survey conducted by the Victor Adding Machine Co. (Courtesy Victor Adding Machine Co., Chicago, Ill.)

1. Questions to be included were chosen by management and phrased to permit convenient and rapid recording of answers on electric punched cards for sorting and tabulating.

2. Supervision was informed of the project at special supervisory meetings.

3. Employees were informed of the project by supervisors. The employees were told that they would receive the questionnaires through the mail at their home addresses and that the company would appreciate their cooperation in checking the questionnaires and returning them to the tabulating agency (not the plant) in the enclosed self-addressed stamped envelope.
4. Questionnaires were received, results were tabulated, and a report of findings was prepared for management.

In the case of the survey conducted for this company, the procedure described above resulted in a 70 per cent return of the questionnaires, a proportion far exceeding that obtained with many questionnaires sent out without previously notifying the employees that a questionnaire would be received.

A summary of the findings of the survey is not of particular importance to anyone not associated with the particular company and, therefore, is not given here. It is sufficient to say that the company made significant use of the results in planning the future activities of the industrial relations department.

A general employee opinion survey of the type described above can also be used as an instrument for the measurement of over-all morale if a few simple principles developed and described by Harris[8] are followed. Harris identified forty-eight questions that would seem to indicate over-all morale from questions appearing on the Victor Adding Machine Company questionnaire illustrated in part in Fig. 10.1. Ten judges (advanced graduate students and staff members in industrial psychology) were asked to check the response to each question that "most strongly represents a favorable attitude toward the company." Any question on which the independent checks of the ten judges did not show at least 80 per cent agreement was discarded at this point. Forty-six questions met this criterion and it is interesting to note that there was unanimous agreement on forty-two questions.

The questionnaires returned by the employees were then divided into two random halves after the halves had been stratified on the basis of several characteristics such as sex, marital status, and length of service. Item analysis against a criterion of total score was then conducted with one of the stratified halves. This resulted in the identification of thirty-six questions, which constituted the over-all morale scale. The reliability of this scale determined by correlating odd *vs.* even scores on the half of the papers not used in the item analysis was found to be .84.

---

[8] F. J. Harris, "The Quantification of an Industrial Employee Survey. I Method," *Journal of Applied Psychology*, 1949, 33, 103–111, also "The Quantification of an Industrial Employee Survey, II. Application," *Journal of Applied Psychology*, 1949, 33, 112–113.

Some typical questions making up the final scale, with the high-morale answer checked, are:

What does your family think of this Company?

Good place to work __________x__________
No opinion __________
Poor place to work __________

How do you like your present job?

Very much __________x__________
Pretty good __________
Not so good __________
Don't like it __________

The measurement of over-all morale by means of questions of this sort extracted from a general employee opinion survey presents some interesting possibilities.

"Theme" Analysis of Employee Communications. A recent and very promising method of measuring the reaction of employees to various aspects of their jobs has been developed by the Employee Research Section of the General Motors Corporation.[9] This work made use of 174,854 letters entered in the General Motors MJC (My Job and Why I Like It) Contest. Letters were entered by 58.8 per cent of all eligible employees. Aside from the evaluation of letters for awarding prizes, the letters proved to be a source of valuable information regarding opinions of employees in the various divisions. To "mine" this source of information, a *theme analysis* of the letters was made. Some of the letters were reviewed for content to identify the "themes" that were being mentioned. Fifty-eight themes were identified. Some examples are: supervision, wages, pride in product, security, employee relations, and pension plans.

After identification of the themes, each letter was analyzed in order to determine what theme or themes it mentioned. An example will illustrate how this was done. Below are excerpts from one letter, with the corresponding theme for each of three portions of the letter.

---

[9] C. E. Evans and L. N. Laseau, "My Job Contest," *Personnel Psychology Monograph No. 1* (Washington, D.C., 1950).

EXCERPTS FROM LETTER	THEME
My job with General Motors has given me what every man wants—*security for his family*	*Security theme*
In less than a month I was getting top rate	*Success theme*
Knowing that a *chance for advancement is* possible	*Opportunity for advancement theme*

Although the participants would be expected to mention only the *positive* factors of their jobs, it was believed that their *lack of mention* of certain factors might be meaningful. Among the various analyses that were made, one consisted of a rank order of the frequency of mention of themes in the letters from the various divisions. A comparison was then made of the rank order of frequency of themes mentioned in *each* division with the rank order for *all* divisions. An illustration of this comparison is given in Table 10.3. This table shows the twenty themes that were mentioned most frequently for all divisions, giving their rank order of frequency of mention. The table also shows the rank order of those same twenty themes from one particular division (No. 48). Although the first six themes as mentioned by employees in Division 48 had the same rank order of frequency of mention as in the corporation as a whole, it can be seen that the rank orders of some of the subsequent themes differed, some of them rather markedly. It can be seen, for example, that *opportunity for advancement* was ranked 14th by all divisions collectively, but was ranked 8th by employees in Division 48; this suggests that the employees in that division perceive their opportunities for advancement to be considerably better than do employees in the company at large. On the other hand, the *medical facilities* in that division are perceived as being below average, as indicated by the fact that for all divisions this theme was 15th whereas for Division 14 it ranked 23rd.

This General Motors approach to theme analysis of employee communications presents some very promising possibilities for the measurement of employee reaction toward various aspects of their jobs.

TABLE 10.3

Rank Order of Selected Themes from General Motors "My Job Contest" for All Divisions and for Division 48

Theme name	*Rank order*	
	All divisions	*Division 48*
Supervision	1	1
Associates	2	2
Wages	3	3
Work type	4	4
Pride in company	5	5
Management	6	6
Insurance	7	9
Pride in product	8	11
Benefits from wages	9	13
Pride in stability of company	10	12
Safety	11	16
Security	12	10
Training, education, experience	13	7
Opportunity for advancement	14	8
Medical facilities	15	23
Teamwork	16	14
Tools, methods, equipment	17	17
Paid vacation	18	20
Cleanliness	19	24
Pride in important work	20	15

Source: Adapted from Evans and Laseau, *op. cit.*

Interviews with Employees. On occasions, the interview is used as a method of obtaining information about employee attitudes and opinions. When this is done, either all the employees, or a sample, are interviewed, being given assurance that the information furnished will not be used for any administrative purpose. To provide such assurance, the interviews usually are conducted by representatives of some outside organization, such as a consulting firm or a university.

Such interviews may be *guided* or *unguided.* In the guided interviews, the interviewer asks a series of questions, each of which may be answered by a simple "yes," or "no," or by some other word or phrase. In the unguided interview, the interviewer asks more general questions to encourage the employee to express himself in whatever manner he wishes. The comments of the employee are used by the interviewer as the basis for assessing the employee's attitudes and opinions.

The guided interview is something like an orally administered questionnaire, the results of which usually can be summarized statistically. The results of the unguided interview, on the other hand, do not lend themselves as readily to quantification. With this method the interviewer usually "rates" the attitudes of employees. In this connection there is some evidence to indicate[10] that such ratings tend to overestimate or underestimate the strength of employee attitudes, and are highly inconsistent.

Aside from such limitations of the interview, it must also be recognized that the cost of this method is high. Except under unusual circumstances, therefore, a questionnaire usually provides a more adequate and practical method of measuring employee attitudes and opinions.

The Exit Interview. Another method of determining both general employee attitude toward the company and also feeling or opinion with regard to specific practices, departments, or supervisors, is provided by the exit interview.

Onarheim,[11] summarizing the values of the exit interview, points out that it helps to retain desirable employees and provides a check on policies covering employment, placement, training, wages, reasons for terminations in different departments, and effectiveness of grievance procedure. Drake[12] also stresses the value of the exit interview in identifying sources of employee dissatisfaction.

An employee who is working for the company often is reluctant in expressing his true feelings. Also, many managements feel that it would not be good industrial relations practice to encourage present employees to give verbal expression to their feelings. But the employee who has terminated his employment usually is quite willing to say what he thinks about the company, and management has little to lose (and often much to gain) by listening attentively to what he has to say. This fact has resulted in an increasing use of the exit interview—an interview conducted with the employee leaving the company. Waters[13] has emphasized that skillfully conducted exit interviews will reveal the cause of high turnover. To obtain the greatest value from the exit interview Hardt[14] urges that the personnel man should conduct the interview outside his office and should take all the time needed to allow the terminating employee to tell his full story in his own words. Any single employee may have grievances for which there is no real basis in plant practices, yet if an appreciable number of employees terminating their employment mention the same situations or practices as unsatisfactory, it is usually safe to conclude that there is a real reason for their discontent. Under such circumstances, it is also a reasonable assumption that employees still on the job are not entirely satisfied with these practices or policies. Information obtained in a series of exit interviews should not, of course, be automatically followed by changes in managerial policy whenever a source of grievance is revealed, but most managements now believe that such information should at least be considered when making future managerial decisions.

---

[10] "Comparison of Scaled Questionnaire and Interview Methods in Appraising Industrial Attitudes," *Research Bulletin 1*, Bureau of Industrial Psychology, University Extension Division, University of Wisconsin (undated).

[11] J. I. Onarheim, "Exit Interviews Help Us Check Personnel Policies," *Factory Management and Maintenance*, 1944, **102**, 121–122.

[12] C. A. Drake, "The Exit Interview as a Tool of Management," *Personnel*, 1942, **18**, 346–350.

[13] J. F. Waters, "Reducing Labor Turnover," *Personnel Journal*, 1949, **28**, 102–108.

JOB SATISFACTION

As indicated earlier, many companies are genuinely interested in employee attitudes and opinions. Mention has been made of various methods used by management for either measuring employee attitudes and opinions, or for obtaining some impression of them. An understanding of employees' behavior, however, cannot be based entirely on their attitudes and opinions relating to their specific work situations. Moreover, it is helpful, in understanding human behavior in general, to know something about the sets of values by which people live, and about the satisfactions associated with the type of work they do. Job satisfaction, of course, is influenced both by the extent to which the work a person does is intrinsically interesting to him, and by his attitude toward the total work situation, including the company, his supervisor, and his fellow workers.

Job Satisfaction in General. To take something of an over-view of job satisfaction of people, let us review briefly some of the results of a survey by Hoppock[15] carried out at New Hope, Pa. In this survey, 309 people in the community were asked to fill out a questionnaire dealing largely with certain aspects of job satisfaction. Responses to certain questions are given in Table 10.4.

---

[14] E. F. E. Hardt, "The Exit Interview: A New Approach," *Modern Management*, 1949, **9**, 10.

[15] R. Hoppock, *Job Satisfaction* (New York, Harper and Brothers, 1935).

TABLE 10.4

Responses to Certain Questions on Job Satisfaction Survey

Question and Response	*Percentage*
Choose the ONE of the following statements which best tells how well you like your job:	
1. I hate it	2
2. I dislike it	2
3. I don't like it	11
4. I am indifferent to it	9
5. I like it	63
6. I am enthusiastic about it	9
7. I love it	5
Check one of the following to show HOW MUCH OF THE TIME you feel satisfied with your job:	
8. All of the time	41
9. Most of the time	27
10. A good deal of the time	8
11. About half of the time	9
12. Occasionally	5
13. Seldom	5
14. Never	5
If you could have your choice of all the jobs in the world, which would you choose?	
22. Your present job	48
23. Another job in the same occupation	16
24. A job in another occupation	36
Which gives you more satisfaction?	
32. Your job	66
33. The things you do in your spare time	34

Source: Hoppock, *op. cit.*, pp. 250–252.

It can be seen that people differ in the degree to which their jobs are satisfying to them. It might be pointed out, however, that many of the people covered by this survey indicated that they liked their jobs. The degree of job satisfaction is somewhat related to the type of job, as shown in table 10.5. This gives the range, and the mean, of the "job satisfaction indexes" of people in each of various occupational classifications. The job satisfaction index, derived from responses to the questionnaire, can range from 100 (minimum satisfaction) to 700 (maximum satisfaction). Although the mean indexes do differ for the various occupational groups, it can be seen that the ranges for the groups are all fairly wide, indicating that job satisfaction, even within each occupational group, varies considerably from individual to individual.

TABLE 10.5

Job Satisfaction Indexes of Five Occupational Groups

Occupational Group	*No. of cases*	*Range of indexes*	*Mean index*
Unskilled manual	55	100–650	401
Semiskilled	74	125–650	483
Skilled manual and white-collar	84	125–675	510
Subprofessional, business, and minor supervision	32	250–700	548
Professional, managerial, and executive	23	300–700	560

Source: Hoppock, *op. cit.*, p. 225.

Another job satisfaction questionnaire has been developed by Brayfield and Rothe.[16] This questionnaire, developed by a combination of the Thurstone and Likert methods, has the primary characteristics of the Likert type of scale. It provides for indicating the degree of agreement with each of 19 statements such as:

My job is like a hobby to me.
My job is usually interesting enough to keep me from getting bored.
It seems that my friends are more interested in their jobs.
Each day of work seems like it will never end.
I find real enjoyment in my work.

It was reported that attitudes measured by this scale correlated .92 with those obtained by Hoppock's[17] questionnaire.

What People Want from Their Jobs. Job satisfaction, in part, is affected by the degree to which the work situation results in the fulfillment of certain values that individuals possess. People look for different kinds of values in their jobs, some placing greater value on, say, security, others on income, and so forth. Although individuals vary in their sets of values, there sometimes is at least moderate consistency in the values of individuals within specific groups. It should be added, incidentally, that the values are subject to at least moderate change, depending upon economic, social, and other influences. An understanding of the value systems of employees can be very useful to management in developing personnel programs. If, for example, management operates on the philosophy that all that the employees are interested in is money, it might be surprised to find that, despite high wages, the employees vote to go out on a strike over some apparently trivial matter.

---

[16] A. H. Brayfield and H. F. Rothe, "An Index of Job Satisfaction," *Journal of Applied Psychology*, 1951, 35, 307–311.

[17] R. Hoppock, *op. cit.*

TABLE 10.6

Relative Importance of 28 Factors to the Nonselling Employees of a Large Merchandising Organization

	Rank
Receiving help necessary to get results expected by management	1
Being encouraged to offer suggestions and try out better methods	2
Being able to find out whether work is improving	3
Reasonable certainty of being able to get fair hearing and square deal in case of grievance	4
Certainty of promotions going to best qualified employees	5
Encouragement to seek advice in case of real problems	6
Being given information about important plans and results which concern the individual's work	7
Being given reasons for changes which are ordered in work	8
Not being actually hampered in work by superior	9
Getting contradictory or conflicting orders	10
Being given to understand completely the results which are expected in a job	11
Pay—*Assurance of increases when deserved*	12
Being invited to offer suggestions when new plans are being considered	13
Feeling that superior understands all about the difficulty of the individual's job	14
Being given to understand completely the general methods which the superior wants followed	15
Complete definition of duties	16
Not being responsible to too many superiors	17
Knowledge of other jobs in the organization which the individual feels capable of handling and would prefer	18
Knowledge of other jobs preferred, even at same pay	19
Red tape in the organization, preventing best work	20
Pay—*compared to that of other jobs of equal importance in the organization*	21
Pay—*compared to that of similar work in other organizations*	22
Treatment when being employed	23
Knowledge of lines of promotion	24
Value of Mutual Benefit Association	25
Being permitted to make important decisions in work	26
Regularity of amount of work	27
Service of Medical Department	28

Typical of studies on this subject is one by Houser[18] who found that the nonselling employees (including unskilled labor) of a large merchandising organization considered the factors shown in Table 10.6 to be the things they wanted from the company. The order in which these factors are listed in Table 10.6 shows the relative importance attached to the factors by the employees. It is interesting to note that the employees of this company ranked the items concerned with present pay twenty-first and twenty-second in importance.

Though studies of other employees in other plants do not consistently show the financial incentive to be so far down the list as it appears in Houser's study, other studies do almost universally show that nonmonetary factors are of significant importance. Hersey,[19] in a study of union and nonunion employees, found that fourteen factors were considered "most important" by the percentages of employees tabulated in Table 10.7.

TABLE 10.7

Percentage of Union and Nonunion Employees Who Considered the Listed Items as "Most Important" Among Factors Related to Their Jobs

	Per Cent of Union Employees Checking Item as "Most Important"	*Rank of Item*	*Per Cent of Nonunion Employees Checking Item as "Most Important"*	*Rank of Item*
1. Employee stock subscription	5	11.5	2	13.5
2. Voice or share in management	13	9.5	6	11.0
3. Fair adjustment of grievances	80	1.0	24	7.0
4. Chance of promotion	28	6.0	47	3.0
5. Steady employment	65	2.0	93	1.0
6. Medical and dental service	0	13.5	6	11.0
7. Safety	57	3.0	21	9.0
8. Amount of pay	49	4.5	51	2.0
9. Working conditions	49	4.5	45	4.0
10. Hours of work	13	9.5	23	8.0
11. Type of man in charge	18	7.5	38	5.0
12. Methods of pay	0	13.5	2	13.5
13. Insurance systems and pensions	18	7.5	36	6.0
14. Chance to show initiative	5	11.5	6	11.0

---

[18] J. D. Houser, *What People Want from Business* (New York, McGraw-Hill Book Company, Inc., 1938), 29.

[19] R. B. Hersey, "Psychology of Workers," *Personnel Journal*, 1936, **14**, 291–296.

Hersey did not find the financial aspect of the job to be quite so unimportant, in comparison with other factors, as did Houser in his study. Yet Hersey did find two things of importance. First, he found that amount of pay was not considered most important by most employees in either the union or the nonunion group. "Steady employment" was checked by more employees as a most important factor in both groups and, in the union group, the factors of "fair adjustment of grievances" and "safety" also outranked "amount of pay." Second, Hersey found that still other factors, though admittedly of somewhat lesser importance than amount of pay to the majority of employees, were considered of first importance by a sufficiently large percentage of employees to justify definite attention by management.

TABLE 10.8

Ranking of Ten Items in Order of Importance by 325 Factory Workers

Rank	*Factor*
1	Steady work
2	Comfortable working conditions
3	Good working companions
4	Good boss
5	Opportunity for advancement
6	High pay
7	Opportunity to use your ideas
8	Opportunity to learn a job
9	Good hours
10	Easy work

Another investigation dealing with the factors that appeal to factory workers has been reported by Wyatt, Langdon, and Stock.[20] In this study, ten factors were ranked by 325 employees in the order of importance shown in Table 10.8.

In this report, as in those mentioned before, many factors besides those of a monetary nature were considered of real importance by the industrial employees studied.

In one part of the study by Hoppock,[21] eighty persons representing a wide range in age, intelligence, occupations, and earnings were interviewed. From these results Hoppock states that no conclusions are drawn except "that what happened in these cases can happen."

---

[20] S. Wyatt, J. N. Langdon, and F. G. L. Stock, "Fatigue and Boredom in Repetitive Work," *Industrial Health Research Board, Report No. 77* (London, 1937).

[21] R. Hoppock, *op. cit.*

The interview case studies corroborate, however, the general conclusion reached above—that is, that job satisfaction is related to a good many things besides financial return. Some of these factors, as summarized by Hoppock, are "relative status of the individual within the social and economic group with which he identifies himself, relations with superiors and associates on the job, nature of the work, earnings, hours of work, opportunities for advancement, variety, freedom from close supervision, visible results, the satisfaction of doing good work, opportunities for service to others, environment, freedom to live where one chooses, responsibility, vacations, excitement, opportunity for self-expression, competition, religion, opportunity for or necessity of traveling, fatigue, appreciation of criticism, security, and ability to adjust oneself to unpleasant circumstances."

Further evidence of the importance of nonmonetary factors in influencing employee morale comes from a study reported by Roberts[22] who made an audit of the morale of 2,500 employees. His findings show the importance to morale of such factors as promotion procedures, favoritism, and the effective informing of employees of company policies. A study reported by Blum and Russ[23] also showed that certain nonmonetary factors—job security and advancement—ranked above pay in their influence upon employee morale.

The fact that neither management nor union leaders necessarily agree with rank-and-file employees on what factors are most important for high morale is shown in a study by Raube.[24] Raube submitted a list of seventy-one morale factors to employees of six companies, fifty executives of these companies, and forty-two labor leaders. Each employee was asked to identify in order of importance the five factors that were most important to him. The executives were asked to predict the rankings that would be assigned to these factors by the employees of their respective companies, and the labor leaders indicated the rankings that they believed represented the relative importance of the factors to their union members. Definite differences in the rankings were found for the three groups. The employees attached most importance to the following five:

---

[22] E. B. Roberts, "Tests to Determine Objectively the Effectiveness of an Industrial Relations Program," *Office Management Series, American Management Association*, 1938, No. 84, 32–37.

[23] M. L. Blum and J. J. Russ, "A Study of Employee Attitude Toward Various Incentives," *Personnel*, 1942, **19**, 438–444.

[24] S. A. Raube, "Factors Affecting Employee Morale," *Studies in Personnel Policy, No. 85* (New York National Industrial Conference Board, 1947).

1. Job security.
2. Compensation.
3. Opportunities for advancement.
4. Employee financial benefits (insurance, and so forth).
5. Informing employees about their job status.

Executive predictions agreed with the employee's ranking in only three of the most important five (job security, compensation, and opportunities for advancement), while labor leaders agreed with the employees in only two (job security and compensation). While compensation and job security were selected as two of the five most important factors by all three groups, there was little agreement in identifying the other factors considered by the employees to be most important.

Another investigation dealing with what men want from their jobs was made by Jurgensen,[25] who asked 3,345 male applicants for employment to rank ten factors in order of importance to them. The ten factors, in the average rank order of importance, are listed below:

1. Job security.
2. Opportunity for advancement.
3. Type of work.
4. Company.

5.5 (a) Pay
5.5 (b) Co-workers } These two tied for 5th and 6th place.

7. Supervisor.
8. Working hours.
9. Working conditions.
10. Benefits.

Here, again, several other factors are considered more important than the amount of pay.

Jurgensen has summarized in convenient tabular form results on this subject reported by several investigators.[26] These results are reproduced in Table 10.9.

---

[25] C. E. Jurgensen, "What Job Applicants Look For in a Company," *Personnel Psychology*, 1948, 1, 433–445.

[26] C. E. Jurgensen, "Selected Factors Which Influence Job Preferences," *Journal of Applied Psychology*, 1947, 31, 553–564.

TABLE 10.9

Rank Order of Job Preference Factors as Obtained by Various Investigators

Source:	Chant*		Wyatt, Langdon, and Stock†	Berdie‡	Blum and Russ§	
Group:	Misc. Workers	Dept. Store Workers	Women Factory Workers	Male H.S. Graduates	Male	Female
Number:	150	100	325	150	181	105
Opportunity for advancement	1	1	5	2	1	1
Steady work	2	2	1	1	2	2
Opportunity to use your ideas	3	3	7	4		
Opportunity to learn a job	4	4	8	7		
Opportunity to be of public service	5	7		8		
Good boss	6	5	4	9	4	3
High pay	7	6	6	3	3	4
Good working companions	8	8	3	5		
Comfortable working conditions	9	9	2	11		
Clean work	10	11		10		
Good hours	11	10	9	6	5	5
Easy work	12	12	10	12		

* S. M. F. Chant, "Measuring the Factors that Make a Job Interesting," *Personnel Journal*, 1932, **11**, 1–4.

† Wyatt, Langdon, and Stock, *op. cit.*

‡ R. F. Berdie, "Can Factors in Vocational Choice be Weighted?" *Occupations*, 1943, **22**, 43–46.

§ Blum and Russ, *op. cit.*

These studies should not be interpreted by management as meaning that the average employee has no interest in his pay check or that he wants to be "fathered" or "mothered" or in any other way made the beneficiary of a paternalistic system. He wants the opportunity to do a job for a reasonable wage, to have an impartial hearing if he thinks he has been treated unfairly, and to be reasonably sure of holding his job as long as his work is satisfactory. In one investigation,[27] the last-mentioned point, job security, was found to be more important than any other factor.

---

[27] D. McGregor, "The Attitudes of Workers Toward Layoff Policy," *Journal of Abnormal and Social Psychology*, 1939, **34**, 179–199.

Added to these factors should be an opportunity for the employee to work with a group that "accepts" him and of which he feels himself to be an integral part. A series of investigations conducted by Mayo[28] and his collaborators in a variety of plants has shown the influence of social and group factors upon the worker's production, quality, absenteeism, and related characteristics. These studies show, among other things, that many of the factors (such as illumination) previously considered to be specific *determiners* of employee productivity operate in a much less specific and tangible fashion than has previously been supposed. For example, in a series of studies[29] of a small group of employees at the Hawthorne plant of the Western Electric Company, an effort was made to determine the effect upon production of varying the illumination. When the lighting was increased, production increased; when the lighting was left the same, production also increased; and when the lighting was *decreased*, production still increased. In this particular experiment, whatever was done had a favorable effect on employee job performance. The only reasonable explanation seemed to be that the employees in the experimental group considered themselves a rather special group of employees, and that they were responding in a favorable manner to whatever changes management made in their working conditions. There is little reason to believe that such factors as a new job evaluation plan, a new incentive plan, a new labor contract, or the careful selection and training of new employees would have entirely taken the place of this *group feeling* in improving job performance. But factors affecting employee *morale* might quickly undermine this group feeling; these factors, therefore, must be considered by management if employee morale is to be kept on the high level that efficient plant operation requires.

---

[28] Several of these studies have been summarized by Elton Mayo, *The Social Problems of an Industrial Civilization* (New York, Andover Press, 1945).

[29] Reported by T. N. Whitehead, *The Industrial Worker* (Cambridge, Mass., Harvard University Press, 1938).

11

ATTITUDES AND HUMAN BEHAVIOR

The last chapter dealt with the measurement of attitudes and morale, and included summaries of the results of several attitude and opinion surveys. These illustrations point up the fact that *individuals* within a given work situation differ in their attitudes and opinions. Aside from obvious individual differences, however, these, and other surveys, quite clearly show differences in attitudes and opinions of *groups* of people in various work situations. Such *group* differences tie in with the implications of earlier statements to the effect that the attitudes and opinions of people are influenced in part by variables in the work situation.

It is then in order to ask *what* variables in the work situation contribute to the formation of employee attitudes and opinions. If one can gain some realistic answers to this question, it might then be possible to attempt to create those conditions which presumably contribute to favorable attitudes and opinions. We will return to this matter in the next chapter; let us for the moment simply indicate that there is strong reason to believe that attitudes and opinions are influenced by a complex combination of factors that is sometimes called "human relations."

Before dealing with this subject, however, we should first consider a point that was raised, but not discussed, in the last chapter—the assumption that frequently is made that employees with favorable attitudes and opinions are, in general, "better" employees than those with unfavorable attitudes and opinions.

EMPLOYEE ATTITUDES AND BEHAVIOR

In a comprehensive survey of studies relating to employee attitudes, Brayfield and Crockett[1] were especially interested in crystallizing, if they could, any general implications of the many studies regarding the relationship between general employee attitudes and "performance." Various aspects of employee "performance" were considered, including the following: job performance as such, absences, tardiness, and tenure (referred to as "employment stability").

---

[1] A. H. Brayfield and W. H. Crockett, "Employee Attitudes and Performance," *Psychological Bulletin*, 1955, 52, 396–428.

Employee Attitudes and Job Performance. It must be stated that the survey brought out information that is somewhat at odds with common beliefs or assumptions. In particular there seemed to be no consistent pattern of relationship between attitudes and opinions on the one hand and actual job performance on the other hand.

Two or three studies will be discussed in order to illustrate the divergencies of the results of such studies. In one study Mossin[2] investigated the sales performance of 94 female salesclerks in a department store in relation to the salesclerks' scores on a job satisfaction questionnaire. Sales performance was quantified by ratings. The correlation between job satisfaction scores and ratings on sales performance was —.03, indicating the absence of any significant relationship.

Another study with somewhat similar results was reported by Bernberg[3] in an aircraft plant. He used four attitude measures that need not be described in detail; briefly, they measured "group morale," the employee's "acceptance" of the formal organization, attitude toward supervisor, and "self-rating" scale. These were correlated with job performance ratings, the correlations ranging from .02 to .05, which are almost as close to zero as correlations can get.

An illustration of a study with different results is one that was carried out at The Prudential Life Insurance Company by The Survey Research Center of the University of Michigan.[4] In this study it was possible to identify certain "high" producing sections and other "low" producing sections, since various sections were doing identical types of office work, and since there were records available of the actual clerical time spent in completing a given amount of work. In one phase of the study four attitudinal variables were developed, namely: (a) pride in work; (b) intrinsic job satisfaction; (c) involvement in the company; and (d) financial and job status. Attitudes on these four variables were obtained by the use of interviews. After deriving attitude scores on these variables from the interview, it was possible to compare the attitudes of employees in the "high" and "low" sections.

---

[2] A. C. Mossin, *Selling Performance and Contentment in Relation to School Background*, Bureau of Publications, Teachers College, Columbia University, 1949.

[3] R. E. Bernberg, "Social-Psychological Factors in Industrial Morale: I. The Predictions of Specific Indicators," *Journal of Social Psychology*, 1952, **36**, 73–82.

[4] D. Katz, N. MacCoby, and Nancy C. Morse, *Productivity, Supervision and Morale in an Office Situation*, Survey Research Center, Institute for Social Research, University of Michigan, December 1950.

The results are summarized in Fig. 11.1. This gives, for the "high" and "low" producing sections, the percentages who had high, medium, and low "morale" in each of the four attitude areas. It was found that "pride in work" bore a significant relationship to the work performance of the various sections.

The primary implication of the various rather diverse studies relating to attitudes and job performance is that satisfaction with one's work situation, or with aspects of it, apparently does *not necessarily* provide strong motivation to outstanding job performance. The general inference from the various conflicting studies is that there presumably are certain *kinds* of circumstances in which job satisfaction is related to work performance, and *other* kinds of circumstances where this is *not* the case. So far the variables that account for these differences have not been adequately identified. As pointed out by Wilensky,[5] there may be a number of such variables such as: economic and technological (employment levels, business cycle, and the like); social and cultural background of employees; nature of work incentives; or the nature of the leadership in the organization.

Employee Attitudes and Absenteeism, Tardiness, and Tenure. Although attitudes do not seem consistently to be related to job performance as such, there is somewhat more consistency in the relationship of attitudes to such aspects of behavior as absenteeism, tardiness, and tenure or "employment stability." Even with these criteria the reported studies do not indicate strongly consistent patterns, but the various pertinent studies suggest at least a moderate (though complex) relationship between attitudes and absenteeism, tardiness, and employment stability. A few examples will be mentioned briefly.

---

[5] H. Wilensky, "Human Relations in the Workplace: An Appraisal of Some Recent Research," in *Research in Industrial Human Relations* (Harper & Brothers, 1957).

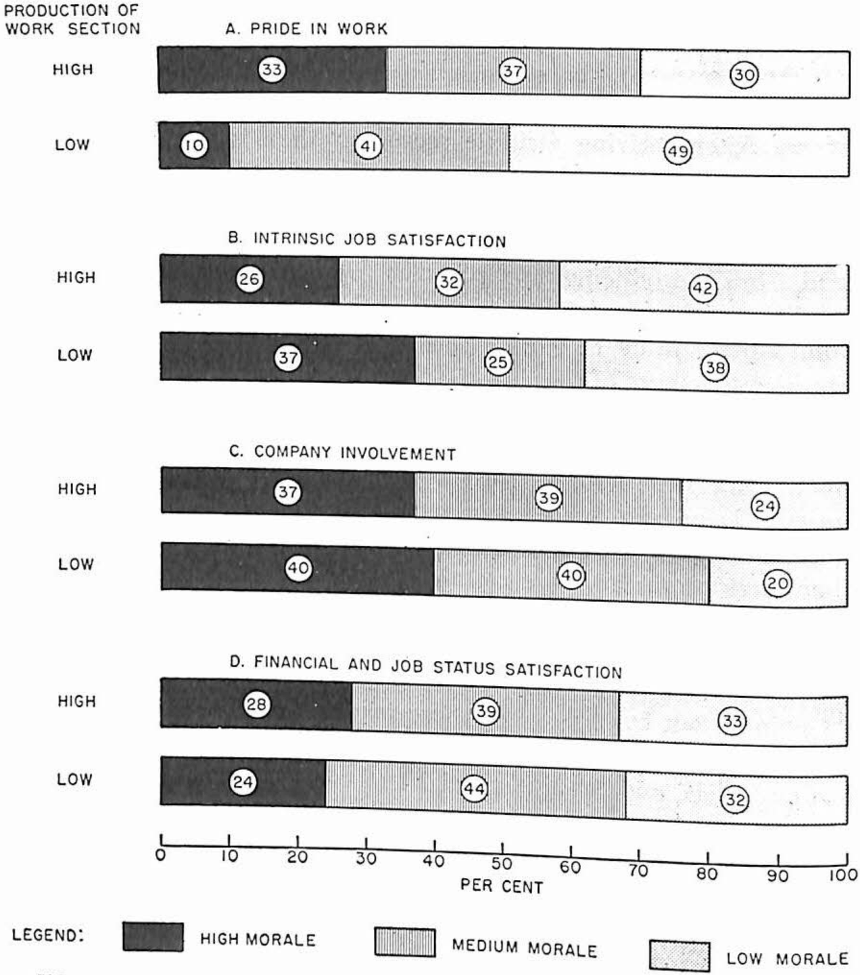


FIG. 11.1. Per cent of employees in high and low producing sections who had high, medium, and low "morale" indexes for each of four morale areas. The relationship between "morale" and section productivity was significant only for the first morale area, namely, "Pride in Work" groups. (Adapted from Katz, MacCoby, and Morse, op. cit., Tables 23, 25, 26, and 32.)

In a study that dealt primarily with attitudes toward merit rating, Van Zelst and Kerr[6] also obtained some information pertinent to our present subject. A total of 340 employees in 14 companies completed a questionnaire that included their own reports of absenteeism and tardiness. In addition, the questionnaire included two questions regarding job satisfaction. Responses to these two questions were combined to give a single index. This job satisfaction index had a correlation with a reported favorable absentee record of .31, and with a reported favorable tardiness record of .26. Both correlations were statistically significant.

---

[6] R. H. Van Zelst and W. A. Kerr, "Workers' Attitudes Toward Merit Rating," *Personnel Psychology*, 1953, 6, 159–172.

In a study by Giese and Ruter,[7] an analysis was made of the average "morale index" of employees in 25 different departments of a mail-order house in relation to tardiness and absence indexes of those departments. The average morale index had no significant relationship with tardiness, but had a statistically significant correlation with absences of —.47; (this can be interpreted as meaning that departments with *most* absences tended to have *low* morale, and vice versa).

In still another absenteeism study, Mann and Baumgartel[8] took a look at various aspects of attitudes and opinions of both white-collar and blue-collar male employees of an electric power company. Later mention will be made of some of the specific results of this study, but in one aspect of it, "over-all satisfaction" of white-collar employees was obtained by the use of a questionnaire. This index was then compared with absenteeism rates for the various work groups (departments, units, and so forth). The results (shown in Fig. 11.2) show that, of the employees in groups with absence indexes of 4 or more, only 22 per cent were "satisfied" with their jobs. By comparison, in groups with absence indexes of 1, and of 2 or 3, the percentages of employees who were "satisfied" were 62 and 52,[9] respectively.

In connection with tenure (or its corollary of turnover) and employee attitudes, Weitz and Nuckols[10] report a statistically significant

---

[7] W. J. Giese and W. H. Ruter, "An Objective Analysis of Morale," *Journal of Applied Psychology*, 1949, 33, 421–427.

[8] F. Mann and H. Baumgartel, *Absences and Employee Attitudes in an Electric Power Company*, Survey Research Center, University of Michigan, December, 1952.

[9] The 52 per cent is an average of the percentages shown in the original data for employees in groups with indexes of 2 and 3; those percentages were 49 and 56, respectively.

[10] J. Weitz and R. C. Nuckols, "The Validity of Direct and Indirect Questions in Measuring Job Satisfaction," *Personnel Psychology*, 1953, **5**, 487–494.

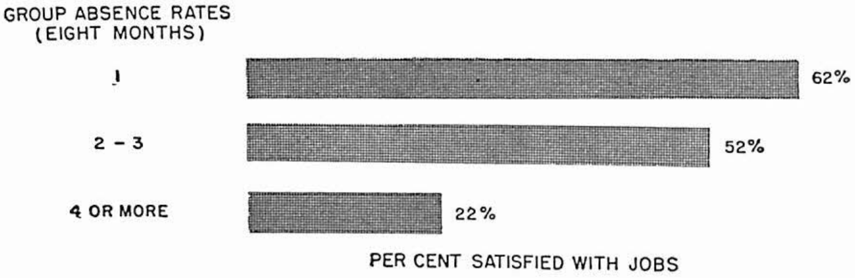


FIG. 11.2. Relationship between over-all satisfaction with company and job, and absence rate, of male white-collar employees. (Adapted from Mann and Baumgartel, op. cit., p. 22.)

correlation of .20 between attitudes of 1,200 insurance agents and their tenure. Kerr[11] reports a significant correlation of .25 for a sample of 98 wage-earners between their "job satisfaction" scores and their self-reported tenure (number of years in the labor market divided by number of employers). In turn, Giese and Ruter[12] report a correlation of "morale scores" with a turnover criterion of —.42 for 25 departments. This *negative* correlation, based on *turnover* indexes, of course indicates a *positive* relationship with *tenure.*

Although some other studies have revealed negligible relationship between employee attitudes and tenure, there nonetheless are a number of situations in which there is some relationship between attitudes and tenure, although the relationship is not very great. This also seems to be true of absenteeism and tardiness.

Comments. In general, then, we cannot say unequivocally that employees with favorable general attitudes are necessarily more *productive* than those with unfavorable attitudes, although there are some situations where this has been found to be the case. There does seem to be some support for saying that attitudes are moderately related to other aspects of employee behavior such as absenteeism, tardiness, and tenure. To maintain an objective, scientific attitude toward this matter, it must be pointed out that the *fact* of this relationship does not *necessarily* imply *cause and effect.* It may be, for example, that those employees who, generally speaking, tend to have favorable attitudes toward their employment situation also are the kinds of people who tend to come to work more regularly, to be on time, and to stay longer on their jobs. In other words, from the strictly scientific point of view, one *cannot* say *if* a company is able in some way to change the attitudes of its employees, that the employees, as a *consequence*, will improve in those aspects of their behavior. To do this requires that one make a basic assumption, namely that there is a cause-and-effect relationship between employee attitudes and employee behavior (such as absenteeism, tardiness, and tenure, and possibly job performance and other aspects of behavior). This requires that one shift gears from a completely objective slant to one in which we must rely on subjective evaluations, value judgments, hints from experiments, and "logic," with perhaps a dash of rationalization thrown in.

---

[11] W. A. Kerr, "On the Validity and Reliability of the Job Satisfaction Tear Ballot," *Journal of Applied Psychology,* 1948, 32, 275–281.

[12] Giese and Ruter, *op. cit.*

Taking such liberties in assessing the potpourri of information regarding human behavior, especially in an industrial setting, the authors feel that there is some justification for believing that the attitudes of employees do have some effect on employee behavior—or, to put it another way, there will be benefits both to employees and to the organization if the organization is able to create those conditions that are conducive to the formulation of favorable employee attitudes.

It may be, however, that the values to the organization are difficult to identify or measure. There may be values *other* than those having a direct influence on employee productivity (or for that matter, absenteeism, tardiness, and tenure). Brayfield and Crockett,[13] for example, suggest that conditions conducive to job satisfaction may have an effect on the quality of the applicants drawn into the organization, on the quality of job performance, and on the harmony of labor-management relations. There may be various types of "hidden" costs in having dissatisfied employees—costs such as having to have more supervision, greater time devoted to handling grievances and complaints, poor community relations and unmeasurable (but nonetheless real) effects on the total organizational efficiency.

Our discussion of employee attitudes so far has been from the point of view of the employing organization. The point of view of employees themselves also should be emphasized. The total welfare of individuals, and of society as a whole, depends, in part, on the satisfactions that people experience in the various aspects of their lives, including their employment relationships. In recent years there has been increasing acceptance on the part of industry of "social" obligations to the communities and to the nation as a whole.

---

[13] Brayfield and Crockett, *op. cit.*

Thus, in terms of both the immediate interests of industry itself and of long-range human welfare, there seems to be adequate justification for the actions taken by management to create work situations that are conducive to the increase of human satisfactions.

This brings us to the question of the motives that serve as the stimulant to management in the development of "human relations." In this connection, a statement by Worthy[14] is uniquely appropriate:

> I agree. . . that gimmicks and devices employed for purposes of manipulation will soon lose their effectiveness. The important question is management's motives in employing the results of human relations research. If its motives are those of narrow self interest, of finding subtler and smoother ways of bending workers to its will, the effort will be worse than useless for it will widen further the gap between workers and management. But if management's motives are sincerely those of better understanding the problems of people at work, of finding ways for making work a more rewarding experience, of discovering its own shortcomings and means for improvement, management's efforts to apply the findings of human relations research are likely to create positive benefits for all concerned.

UNDERSTANDING HUMAN BEHAVIOR

If we accept the premise that the attitudes, work satisfactions, and behavior of employees are influenced in part by conditions or factors in the work situation, it is then in order to try to ascertain *what* such conditions are, and what their influences are. Any attempt on the part of management to create those conditions which are conducive to favorable changes in employee behavior or attitudes, however, must be predicated on an understanding of human behavior. Before discussing such conditions or factors, therefore, let us review briefly some of the basic aspects of human behavior.

Causation in Human Behavior. Human behavior of any kind is caused. Thus, if the foreman "blows his top" over a trifle, such behavior can be said to be brought about by certain causal influences, although it may be difficult to identify them.

---

[14] J. C. Worthy, "Comments on Mr. Wilensky's Chapter," in *Research in Industrial Human Relations, op. cit.*

In the jargon of the theoretical and experimental psychologists, a *stimulus* acts upon the *organism* to produce *behavior* (a response). A simple illustration of this is the instance of a person touching a hot stove. The stimulus (the hot stove) acts upon the organism (the person) to produce behavior (to remove the hand). Most of the situations in everyday life, however, are much more complex than this. For example, when driving in traffic the driver has many "stimuli" in his environment, such as other cars, pedestrians, the traffic lights, the road and road conditions, and perhaps a back-seat driver. His driving behavior is the consequence of an *interaction* of the stimuli and the driver. The driver himself has had previous driving experiences which will influence his behavior in a given situation. Further, his *perception* of the stimuli may influence his behavior. For example, if the driver perceives the speed of a car coming toward him as being less than it actually is, he will behave as though it were travelling at the slower speed; for example, he may then attempt to pass a car in front of him.

In a somewhat parallel vein, a clerk in the complaint department of a store may look at the long line of customers, realize that the closing time is coming (before which she cannot possibly handle all the customers), and may "perceive" their complaints to be directed toward her personally rather than toward the store. Her resulting behavior may be to act in a very snippy and discourteous way to the customers, which, of course, may cause them to take out their wrath directly on her, thereby accentuating the whole situation.

In this process of responding to stimuli, it is very easy to see how the previous experiences of the individual and the attitudes of the individual can affect the resulting behavior. Previous experiences, especially those which have been repeated often enough to form behavior habits, can have the effect of causing a person to persist in a particular type of behavior in situations of a given type. Attitudes, in turn, can affect our perceptions of events or objects in our environment and can provide a predisposition to react in a particular manner.

Human Needs and Incentives. In the above discussion it has been indicated that various individuals behave differently in the same situation, these different *behaviors* being the consequences of the *individuals,* since the *stimuli* are the same in a given situation. In large part these differences among individuals can be attributed to differences in their "motivation." To understand some of the principles of human motivation, however, we must consider two basic aspects of motivation, namely *needs* and *incentives.*

Needs are the felt wants of individuals. They are sometimes also referred to as *drives* or *desires. Incentives* (also referred to as *goals*) are the objects or behaviors that satisfy the needs. Thus, *thirst* is a *need,* and *water* is an *incentive.* Water (the incentive) satisfies thirst (the need).

There are different types of needs, as indicated below:[15]

Innate Needs. These are primarily physiological in nature, and include hunger, thirst, elimination, sex, and temperature regulation. Some innate needs presumably do not have a strictly physiological basis. Smith,[16] for example, suggests that "safety" is a need that is intrinsic to human and animals.

Acquired needs. Throughout life, particularly during the formative years, people acquire some needs by reason of their experiences. Acquired needs, then, are *learned* through our experiences. Thus, people living in the United States today have "needs" that years ago never were even conceived as possibilities. Through succeeding generations, "luxuries" to one generation become "necessities" to the next. Technological developments and economic progress have made many goods and services available as "incentives." Contact between people, and present-day means of communication such as magazines, newspapers, radio, and TV, have contributed to our development of needs that such products and services can fulfill.

Social needs. Special comment might be made about certain types of needs that can be called "social" needs. These include the need for self-achievement, for "belonging" or for being with others, and for "status." Although the social needs probably have certain innate aspects, they are markedly influenced by the cultural environment.

Interaction among needs. Human motivation is not quite as simple as might be implied by considering needs as falling into the above three classes. In actual life the needs of people become intricately intertwined. Thus, a person who has a strong need for status may eat in the "best" restaurants. This he does, not because he is hungry, but because he feels his "status" will be increased by being seen in the "right" places.

---

[15] N. R. F. Maier, *Psychology in Industry* (2nd ed.; Houghton-Mifflin Company, 1955).

[16] H. C. Smith, *The Psychology of Industrial Behavior* (McGraw-Hill Book Company, Inc., 1955).

The interactions of needs are unique for each person. Thus, each person has his own sets of values relating to incentives that are different from those of other people. This, in turn, leads to our difficulties in interpreting the actions of other people, since *we* have values different from *theirs,* and therefore their behavior does not make sense from *our* point of view. Many of the misunderstandings in industry can be attributed to the fact that people differ in their sets of values or needs. Management representatives, for example, may not be able to understand the motivation that induces employees to join a union. Or an employee may not understand why the foreman insists on having a particular job done on time.

It should be emphasized that acquired needs can be, and frequently are, much stronger than those of an innate nature—as long as there is reasonable satisfaction for those of an innate nature. Thus, as long as people are reasonably well fed and have water, hunger and thirst have little influence on their behavior. In a condition of deprivation, however, such needs would become distinctly dominant.

Incentives. Incentives can be either positive or negative. A positive incentive is one that attracts a person like a magnet, as, for example, the prospect of a promotion or high take-home pay. On the other hand, there are some consequences or events that people seek to avoid; these are negative incentives. An example is that of the pressure exerted by a work group on its members to restrict production. Thus, an employee might have the positive incentive to produce in order to increase his piece-rate earnings, but there is the negative incentive of group disapproval if he does so. Here the individual's sets of values will dictate whether he will, or will not, resist the group pressures.

One other point should be made about incentives, this being the distinction between *real* and *substitute* incentives. Given a particular need, there typically is a certain incentive that most logically would fulfill that need. Frequently, however, it is not possible to achieve that particular incentive, or its costs (in terms of other satisfactions) might be excessive. Thus, a person may visualize some *other* incentive as being a substitute for the "real" one. A man may fail to achieve a particular goal, such as becoming a doctor, and redirect his energies toward becoming a medical technician. Or an employee may have his heart set on becoming a foreman, but when this goal fails to materialize, he may become more active in his lodge and ultimately end up as its president, thus fulfilling his need for status.

Frustration. The failure or inability to achieve one's goals may bring about frustration. Such frustration, in turn, can be the root of various types of behavior. In a frustrating situation one person might then set his sights on a substitute incentive, as mentioned above. Such an individual, failing to achieve one goal, settles for another one that may largely fulfill his need. Generally speaking this is a fairly satisfactory method of adjusting to the situation. Some persons, when facing a frustrating situation, adopt a problem-solving approach in attempting to achieve their goals. The accomplishments of many famous people throughout history probably are reflections of such stimulated motivation. Where the goals are achieved through such problem-solving, the frustrations are overcome adequately.

The adjustments of some people to some frustrating situations, however, are less desirable. Maier[17] lists four different undesirable types of frustrated behavior: aggression, regression, fixation, and resignation.

Aggression is a form of behavior involving some type of attack toward another person or object. The 4-year-old boy who beats his playmate over the head with a baseball bat is obviously exhibiting aggressive behavior. The forms of aggression of some adults are much more disguised and subtle. A tidbit of gossip, or a casual question over a tea cup or at the coffee bar, can be as devastating to the victim as the baseball bat.

Regression is the tendency for a person in a frustrating situation to revert to some earlier form of behavior, typically that of children. The temper tantrums, crying jags, emotional outbursts, and pouting of adults are manifestations of regression.

Fixation is the tendency of a person to persist in a particular kind of useless behavior, even though he knows the behavior is useless. An example is that of continuing to telephone an individual when the caller keeps getting no response and knows the person in question is out. The tendencies of some people in industry to resist change probably are indications of fixation behavior.

---

[17] N. R. F. Maier, *op. cit.*

Resignation is the tendency of frustrated people to give up. They take the point of view that there is no use trying to put forth any extra effort. This depressed, negativistic attitude is characteristic of employees who, for example, keep telling others (and themselves) that the cards are stacked against them and that it is futile to do anything about it.

Frustration is part and parcel of human life. The methods individuals adopt for dealing with frustration, however, determine whether the frustrations are the cause of *undesirable* behavior as opposed to serving as the impetus for *desirable* behavior.

Individuals, then, have certain needs which serve as drives to motivate their behavior. Incentives are the specific goals toward which the behavior is directed. In attempting to achieve such goals, however, frustrations frequently occur; these, in turn, can be met with adequate adjustments or with inadequate adjustments. Considering the facts that individuals vary in their "needs," in their perceptions of the incentives that would fulfill those needs, and in the manner of dealing with the frustrations that block the achievement of such incentives, it is not surprising that the foreman sometimes just can't figure out why "Joe behaves the way he does." The important points for the foreman to keep in mind are that Joe's behavior (whatever it may be) is caused by the interaction of the factors discussed above, and that he (the foreman) needs to understand what makes Joe "tick" in order to stand a chance of appreciably modifying Joe's behavior or attitudes.

12

HUMAN RELATIONS

The development of good human relations in industry and business depends essentially on the creation of those circumstances or conditions within which affirmative attitudes and motivations may be expected to develop. Such a program must have two essential features. In the first place, there probably are certain circumstances or conditions that generally speaking can foster affirmative attitudes and motivations with *groups* of people; to the extent that this is so, appropriate actions by management can have a broad, desirable effect. In the second place, the fact of wide individual differences in motivation suggests that *individuals* need to be treated differently, at least to some extent. This bears on the point previously made that a supervisor needs to know what makes an *individual* "tick" in order to take actions that might influence that person's actions or attitudes. Let us first consider some of the factors that seem to form the basis for certain general principles of human relations. In doing so, we will be particularly interested in considering those factors which, on the basis of experimental evidence, have fairly clearly been found to have some bearing on human relations.

SUPERVISION

The supervisor is generally recognized as a keystone in the development of sound human relations. To his subordinates he represents the organization. His personality and behavior typically influence the attitudes and behavior of those whom he supervises. Their reaction to the organization will be flavored by their reaction to him—be it bad or good.

Supervision and Employee Performance. Let us first review briefly some of the evidence that points to the importance of the supervisor in human relations. As one example, further reference will be made to the study by Katz, MacCoby, and Morse[1] in the Prudential Life Insurance Company. It will be recalled that the various sections were characterized as "high" producing or "low" producing sections on the basis of available production records. All of the supervisors of both the "high" and "low" sections were interviewed to determine: (1) a description of supervisory behavior, and (2) attitudes toward own job, subordinates, superiors, the company, and company policies. It was then possible to determine what relationship there was, if any, between supervisory practices and attitudes on the one hand, and section productivity on the other hand.

On the basis of the interviews the supervisors were classified as "employee centered" or "production centered." To help to characterize this distinction, following are the verbatim answers of two supervisors to a question regarding the most important part of the job:

> Supervisor X: "Well, there are two things—keeping the section running smoothly; keeping the clerks happy; keeping production up; making impartial assignments of work; making the proper decisions on some difficult cases involving some payments and maybe a premium that hadn't been paid before." (This supervisor was classified as "employee centered.")
>
> Supervisor Y: "Well, the most important part is to get the reports out. The biggest thing is to get the work out." (This supervisor was classified as "production centered.")

In a similar manner the supervisors were classified as "democratic" *vs.* "authoritarian," and as having "high or average judgment" *vs.* having "poor judgment." The results of comparisons on these classifications for the supervisors of the "high" and "low" sections are shown in Fig. 12.1. This shows that the supervisors of the "high" sections tended to be "employee centered," "democratic," and to have "high or average judgment." The supervisors of the "low" sections, in turn, tended more to be "production centered," "authoritarian," and to have "poor judgment." Although it cannot positively be said that these and other associated characteristics were specifically the *cause* of the sections being high or low in production, the weight of logic and inference tend to support this conclusion.

---

[1] D. Katz, N. MacCoby, and N. C. Morse, *Productivity, Supervision, and Morale in an Office Situation,* Survey Research Center, Institute for Social Research, University of Michigan, December 1950.

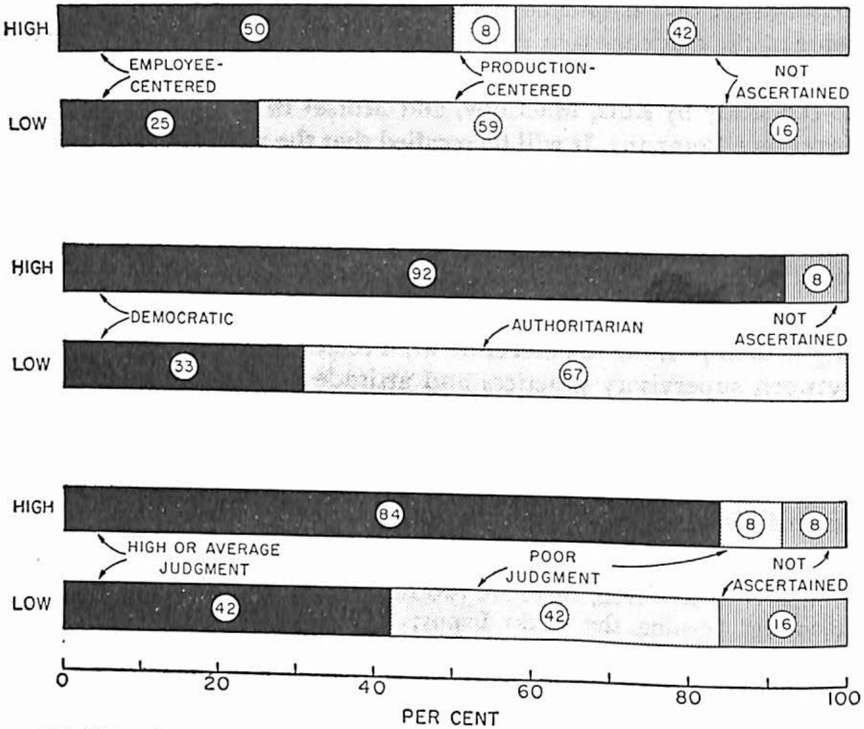


FIG. 12.1. Supervisor characteristics in relation to "high" and "low" producing work sections in insurance company. (Adapted from Katz, MacCoby, and Morse, op. cit.)

Another study that points to the influences of the supervisor on the performance of the work group is one reported by Lawshe and Nagle.[2] In one plant of the International Harvester Company 208 employees in 14 departments filled out an attitude questionnaire which included 21 items regarding the individual's attitude toward his immediate supervisor. In addition, the 14 departments were rated by six executives on their judgment of work output. This was done by the paired comparison method. The executives were supplied with paired comparison forms and instructed to indicate ". . . the department in each pair which is, in your opinion, doing its job better." Although such ratings do not constitute an objective criterion of department performance, the reliability of judgments among the six executives was very high ($r = .88$), suggesting that there were

---

[2] C. H. Lawshe and B. F. Nagle, "Productivity and Attitude Toward Supervisor," *Journal of Applied Psychology*, 1953, 37, 159–162.

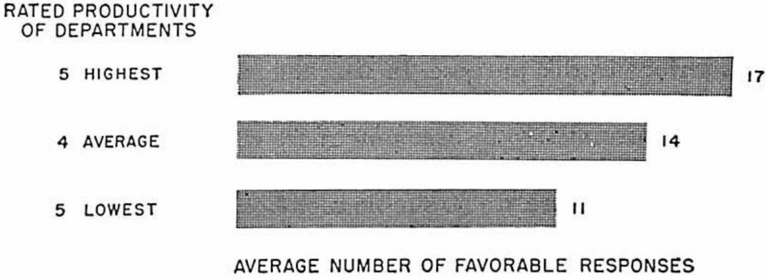


FIG. 12.2. Attitudes toward supervisor (based on responses to attitude questionnaire) of 208 employees in 14 departments of industrial plant in relation to rated productivity of departments. (Adapted from Lawshe and Nagle, op. cit.)

differences in the performance of the various departments that were perceived quite consistently by the executives.

The results of this study are given in Fig. 12.2. Here we see a relationship between attitudes of employees in the various departments toward their respective supervisors and the rated productivity of the departments. In this study the correlation for the 14 departments between average employee attitude (based on the questionnaire) and rated productivity was .86.

From evidence such as the above, there is the strong implication that the supervisor's role is critical in establishing the tone of human relations in the work group, and that the effectiveness with which he does this has a fairly direct bearing on the work performance of the group. Granting this, it is then entirely in order to ask: *What can be done to improve the quality of supervision?* Although there are no final answers to this question, there probably are certain general directions to pursue in attempting to achieve this objective.

Selection of Supervisors. In the first place, if there is reason to believe that the present supervisors in an organization are not the *kind* of people who make good supervisors, management could consider the development of a supervisory selection program. For this purpose, the procedures discussed in Chapter 5 would be appropriate. Fig. 5.1 shows the results of one supervisory selection program.

Training of Supervisors. For both present and new supervisors, appropriate training can aid in the development of supervisory skills. Supervisory training was discussed in Chapter 9, so it need not be discussed further here.

Self-Improvement of Supervisors. A supervisor who wishes to improve his own supervisory skills probably can do so to some degree if he is serious in this purpose. In fact, the recognition by a supervisor that he could become a more effective supervisor is undoubtedly the most important step in *becoming* more effective, when this recognition is coupled with serious intent to do something about it.

Unfortunately, there are no simple cook-book procedures that one can follow to such success. For better or for worse, life is not quite that simple. There are, however, certain general conditions and practices that tend to contribute to the development of improved human relations.

Understanding individuals. A supervisor should develop the facility of understanding individuals—of "reading between the lines" of what they say to recognize the basic needs that motivate them. Although this comes hard for some people, there are no short-cut methods for getting on the same wave length as the employee. In this process, however, a sympathetic ear is essential in listening to employees. Since each person's attitudes and behavior seem, to him, very reasonable, it does violence to a supervisor-employee relationship if, for example, the supervisor treats an employee's comments lightly or with ridicule.

Opportunity for job satisfaction. There must exist within the work situation the reasonable opportunity for employees to achieve at least some degree of job satisfaction. This opportunity is only partially within the hands of the supervisor. The essential character of the work and working area, the conditions of hours and wages, the company policies, and the like, are essentially the responsibility of higher management. The supervisor, however, does have the opportunity to create a desirable social climate within the department, and usually he has the responsibility for assignment of individuals to various jobs within the department.

Dealing effectively with individuals. The supervisor must have the personal qualities and skills necessary to deal effectively with individuals and with the group as a whole. This does not mean that he has to be a silver-tongued orator or a Dale Carnegie, but he does need to be able to get his ideas across and to command the respect and support of those who work for him.

Familiarity with human relations practices. Further, the supervisor should be familiar with, and able to apply, certain principles and practices that have become crystallized through experience and research as being effective in human relations. A few examples will be used to illustrate such principles and practices. In the study reported by Katz, MacCoby, and Morse,[3] for instance, the supervisors were characterized in terms of the "closeness of supervision," meaning the degree to which the supervisor checks up on his employees frequently, gives them detailed and frequent instructions, and, in general, limits the employees' freedom to do the work in their own way. (As some employees might put it, "close" supervision is when the supervisor "breathes down your neck all the time.") The analysis of results of the study showed that over 90 per cent of the "low" production supervisors gave "close" supervision, whereas only 50 per cent of the "high" production supervisors operated in this way.

In a somewhat corresponding study of railroad workers, Katz, MacCoby, Gurin, and Floor[4] found that the supervisors of the "high" producing railroad gangs (as contrasted with supervisors of the "low" producing gangs): (1) spent more time on supervisory duties than on nonsupervisory duties; (2) were perceived by their employees as generally better in planning; and (3) took more personal interest in their employees. There is also evidence from these and other studies that successful supervisors tend to delegate more authority than those who are less successful.

Aside from the human relations practices mentioned above there are certain others that will be mentioned later in further detail. These include the development of good communications, of opportunities for employees to participate in some measure in matters affecting the department, and of cohesive work groups. In all of these the supervisor plays a key role.

Characteristics of Supervisory Behavior. Successful supervisors obviously differ widely in their behavior. Nonetheless, there seem to be certain behavior patterns that are to some degree characteristic of those persons who perform effectively as supervisors. These behavior patterns were investigated in a study by Roach[5] in which a check list of 328 behavior statements was used. These statements had been gleaned from essay descriptions of 35 "good" supervisors and 35

---

[3] Katz, MacCoby, and Morse, *op. cit.*

[4] D. Katz, N. MacCoby, A. Gurin, and L. G. Floor, *Productivity, Supervision and Morale Among Railroad Workers,* Survey Research Center, Institute for Social Research, University of Michigan, 1951.

[5] D. E. Roach, "Factor Analyses of Rated Supervisory Behavior," *Personnel Psychology,* 1956, **9,** 487–498.

"poor" supervisors that had been written by their superiors. The completed check list was then used by the managers for 245 white-collar supervisors, and the results were subjected to a factor analysis which revealed groupings of behaviors that tended to occur in combination. The fourteen factors that were identified by this process are listed below, in a few cases with brief explanatory comments:

1. *Personal Compliance.* (Degree to which supervisor does what is expected of him, such as setting a good example by being on time, and following through on assignments.)
2. *Job knowledge.*
3. *Direction of group performance.* (Characterized by ability to plan and organize, to carry out procedures, and to see that things function smoothly.)
4. *Rewarding performance and thoroughness of employee evaluation.*
5. *Company loyalty.*
6. *Acceptance of responsibility (decision-making).*
7. *Group spirit.*
8. *Personal drive (motivation).*
9. *Impartiality.*
10. *Poise and bearing.*
11. *Consideration.*
12. *Open-mindedness.*
13. *Cheerfulness.*
14. *Approachability.* (Characterized by extent to which he is willing to sit and talk with his employees, and to which they are willing to talk things over with him.)

Although supervisors vary in their patterns of these behavior characteristics, these can be thought of as "dimensions" of desirable supervisory performance.

COMMUNICATIONS

In recent years a great deal has been written and spoken about "communications" in industry. In fact, the development of "communication programs" in industry almost reflects something of a band-wagon or keeping-up-with-the-Joneses effect. There is, however, a rational explanation for this current interest in communications in industry. For this explanation it is necessary to go back into industrial history.

In the years gone by, industrial and business enterprises were typically much smaller than they are today. In a small concern, each employee personally knows the manager, as well as other employees in the concern. This close-knit organizational arrangement permits the easy and informal exchange of information. With the increase in the size of organizations, however, several intermediate layers of management and supervision are inserted between top management and the bulk of the employees. This distance, along with size, simply resulted in the breakdown of the informal, first-hand interchange of yesteryear. This void was not even recognized as a void by some companies for some years. The increased management-labor problems of relatively recent industrial history have served to focus management's attention on the whole area of employee relations. It is probably in connection with this shift that "communications" have come to be recognized as a very essential ingredient in employee relations.

What is Meant by Communications? Communication has been defined as "all of the processes through which information, attitudes, ideas or opinions are transmitted and received, providing a basis for *common understanding and/or agreement*."[6] It has been suggested by Habbe, in turn, that effective communication "creates, or promotes, a better understanding between persons and groups." Both of these expressions place an emphasis on the *effectiveness* of communications rather than simply describing mechanics of transmitting information. The processes of talking to unwilling ears, or passing out pieces of paper, do not insure that the "message" gets through to the person or group for which it is intended.

Within an industrial or business organization communications can flow in at least three directions, namely *down* (from management down through the lower levels of the organization), *up* (from employees to intermediate-level personnel up to higher levels), and *across* (from people on one organizational level to others on the same level). The "content" of communications varies, depending in part on their "direction." For example, downward communications can include instructions, orders, statements of policy and procedures, as well as information that is more general in nature, and upwards communications can include such things as suggestions, attitudes, and grievances.

Methods of Communication. Communications may be either *formal* or *informal*. Formal communications are those that are provided by specific procedures, such as written instructions or policies. Informal communications follow the "grapevine" route, and can cover not only fact but also rumor. Although the informal route can be a very effective method of transmitting information (or misinformation), our primary interest will be with the formal types of communication. These may be either written or oral. A listing of some specific types of communications is given by Habbe,[7] as follows:

---

[6] P. E. Lull, F. E. Funk, and D. T. Piersol, "Business and Industrial Communication from the Viewpoint of the Corporation President," Speech Department, Purdue University, 1954.

1. Attitude surveys.
2. Grapevine.
3. Interviews (casual, counseling, employment, exit, induction, merit rating, open door, training).
4. Mechanical devices (public address system, radio, television, telephone).
5. Meetings (employee-supervisor, group, induction, mass, social-recreation, training).
6. Open house.
7. Suggestion system.
8. Union activities (collective bargaining, grievance procedures, labor-management committees, union meetings).
9. Written communications (bulletin boards, employee booklets, house organs, information racks, letters, posters, reports).

The strengths and weaknesses of some of these media of communication, as based on company experience, are summarized in Table 12.1.

TABLE 12.1

Strengths and Weaknesses of Different Media of Communication

Medium	*Strengths*	*Weaknesses*	*Remarks*
Attitude Survey	Gets ideas up. Potential morale-builder.	Findings may be equivocal. May boomerang if not followed up.	Of considerable value in many situations.
Grapevine	Personal. Of, by, and for employees.	Often inaccurate. Does not reach everyone.	Thrives when other channels of information are closed.
Interview	"Natural." Two-way. Personal. Potential morale-builder.	Hard to train people to be good interviewers.	A basic medium. Can be used in almost any situation for dealing with all kinds of employer-employee relations.

---

[7] S. Habbe, "Communicating with Employees," *Studies in Personnel Policy,* No. 129, National Industrial Conference Board, 1952.

TABLE 12.1 (Cont.)

Medium	*Strengths*	*Weaknesses*	*Remarks*
Public Address System	Fast. Enables management to reach all workers at one time.	Impersonal. One-way. Suitable for short messages only.	Helpful in some matters. Not a basic medium.
Telephone	Convenient. Two-way. Fast.	Not face-to-face.	Indispensable.
Meeting	Face-to-face. Potential morale-builder. Faster than talking individually to employees.	Often one-way. Hard to train people to lead meetings well. Unless done well, audience may regard as waste of time.	Basic and universally used. Can go over big or can flop. The important thing is to learn how to run a meeting.
Open House	Potential morale-builder. Increases information about company.	Gives only superficial view of company's operation.	While arranged especially for employees' families and for residents of the community, this medium can also promote intra-company communication.
Suggestion System	Gets ideas and gripes up to higher management. Voluntary.	Majority of ideas are rejected. Rather slow. Generally used only by minority of employees.	If other media are lacking, a suggestion system may be used by the employees to get gripes, as well as ideas, to management's attention.
Union Activities	Activities often personal. Many deal with matters that are fundamental to employees.	Leadership may not be local. Activities seldom reach all workers.	Value in terms of communication depends on strength of union, leadership, relations with management, and ability to keep close to membership.
Bulletin Board	Official. Convenient.	Impersonal. Only one-way. Suitable for certain types of communications only.	Real but limited function to perform. Nearly all companies use them.
Employee Booklet	Official. Convenient. Comprehensive. Wide coverage.	Impersonal. Only one-way. Needs to be revised rather frequently.	A chief medium for telling employees (new ones especially) about the company.
House Organ	Wide coverage. Report current happenings.	Mostly one-way. Contents often trivial.	General feeling: not the effective communicator it could be.

Source: Adapted from Habbe, *op. cit.*

Various general advantages have been given for written and oral communications, some of these being listed below:[8]

Advantages of written communications:

Authority.
Accuracy.
Permanence.
Coverage (can be of adequate length, and can be duplicated and distributed to all individuals who should receive it).
Miscellaneous (retention rate from reading has been claimed to be high; written communications are economical).

Advantages of oral communications:

Personal.
Two-way (ideas can be exchanged, questions asked, and so forth).
Flexible (can be adapted to receiver; can be presented with "right" tone of voice, with smile, gesture, and so forth).
Effective.
Miscellaneous (easy, simple, fast; places responsibility for communication where it belongs—on immediate supervisor).

Since we are discussing the views of people regarding communications, it would be pertinent to summarize briefly the results of a survey of the viewpoints of 51 corporation presidents regarding communications.[9] The corporations represented were among the large ones in the country. Some of the statistical results are given in Table 12.2; it should be added, however, that some of the questions and responses have been abbreviated in this table.

TABLE 12.2

Summary of Responses of 51 Corporation Presidents to Questionnaire Regarding Business and Industrial Communications

Question and Possible Responses	*Number Giving Response*
1. Preferred method for getting across very important policy (each gave *two* choices)	
Management meeting	44
Personal interview	27
Management bulletin	16
Inter-office memo	14
Telephone	1
2. Do you believe there is a relationship between communication and productivity?	
Definite relationship	48
Slight relationship	2
No relationship	0
(No response given)	1

---

[8] Habbe, *op. cit.*

[9] Lull, Funk, and Piersol, *op. cit.*

TABLE 12.2 (Cont.)

Question and Possible Responses	*Number Giving Response*
3. Importance of oral and written communication	
Oral more important	22
Both equally important	24
Written more important	1
4. Relationship between "breakdowns" of communication and labor disputes and strikes	
Definite relationship	39
Slight relationship	7
No relationship	3
(No response)	2
5. Causes of breakdowns in communication	
Inadequate use of media	34
Lack of communication ability—management	32
Inadequate training program	23
Management withholds information	19
Little opportunity for communication up	16
Lack of communication ability—foremen	15
Union meddling	10
Confusion of authority	8
Clashing personalities	3
Communicator not well liked	0
6. Nature of communicative ability	
Combination of talent and skill	46
Can be learned	35
Personality may prevent development	15
Developed by years of experience	13
Nearly all top management possess	12
Natural gift or talent	5
Related to formal education	0
7. Training for management personnel	
Yes	49
No	0
(No response)	2
8. Training at what levels	
Top management	33
Middle management	44
First-line supervisors	45
9. Factors affecting reception of person's ideas	
Authority of communicator's position	
Very much	37
Somewhat	14
Little	0
Personal regard for communicator	
Very much	36
Somewhat	12
Little	1
(No response)	2

Source: Adapted from Lull, Funk, and Piersol, *op. cit.*

In our present context of communication methods, we can see, from question 1, that there was a strong preference for oral communications (management meetings and personal interviews) for the purpose of transmitting "very important" policies to managerial personnel. In obtaining the data tabulated in question 1 of Table 12.2, it might be added that each company president checked the *two* preferred methods. In choosing the two methods, 21 of the 51 presidents chose two of the oral methods. The majority (31), however, chose a combination of oral and written methods.

For other communication situations, of course, the preferences of executives would be expected to be different. Information about orders received, for example, typically would be written down. (Mention will be made later of some of the other results from this study.) Table 12.1, which summarized the strengths and weaknesses of various media, suggests situations where specific types of media might be used.

The information given above regarding the relative advantages of different types of communications is based largely on opinions of management personnel. An interesting study on the effectiveness of various methods presents some factual evidence to go by. This study, carried out by Dahle,[10] involved the "transmission" of information in various departments of an industrial plant by five different methods. Later, tests of information were given to the employees in order to determine how much of it they had learned. The results, given in Table 12.3, show that a combination of oral and written methods was most effective, as indicated by the fact that employees with whom this method had been used had an average test score of 7.70, as contrasted with an average of 6.17 for oral only, and 4.91 for written only. (The written messages were duplicated materials passed out to the employees.) The bulletin board and grapevine methods resulted in average scores of 3.72 and 3.56, respectively.

The results of this study correspond essentially with the survey of 51 presidents (item 3 of Table 12.2) in suggesting that a combination of oral and written communications, or of oral communications alone, is more effective than using written communications by themselves.

---

[10] T. L. Dahle, "Transmitting Information to Employees: A Study of Five Methods," *Personnel*, 1954, 31, 243–246.

TABLE 12.3

Average Scores on Information Test of Employees with Whom Various Information Methods Had Been Used

Method of Communication	*No. of Employees*	*Average score on Test**
Combined oral and written	102	7.70
Oral only	94	6.17
Written only	109	4.91
Bulletin board	115	3.72
Grapevine only	108	3.56

* All adjacent values in this column differ at the 5% level, except the last two.
Source: Dahle, *op. cit.*

Effects of Communications. Prevailing opinion among executives and personnel administrators supports the contention that good communications are desirable, and that there is some relationship between communications and various criteria of organizational effectiveness. In the survey of 51 executives mentioned above, for example, 48 expressed the opinion that there is a definite relationship between communication and production (item 2 in Table 12.2), and 39 said there was a definite relationship between "breakdown" of communications and labor disputes and strikes (item 4).

It might be pointed out, parenthetically, that the causes of breakdowns in communication were attributed by the corporation presidents primarily to inadequate use of media (34 of the 51), lack of communication ability on the part of management (32), and inadequate training (23), as shown in item 5 of Table 12.2.

It is all well and good for people (including the authors of this text, incidentally) to go around saying that communication is a fine thing—that it is the "lubricant in the wheels of industry"—but the expression of such opinions does not answer the question as to whether good communications are, or are not, of positive value to industry or the employees.

In answering this question let us first review the results of an attitude survey conducted in two plants by the National Industrial Conference Board.[11] In Plant B, special attention had been paid to communications, including the inauguration of regular work-unit meetings in which each foreman met with his employees and discussed matters of common interest. In the other plant (plant A) no such program had been developed. A comparison of the responses of employees in the two plants to a few of the questions provides some revealing contrasts. Results from a few of the questions are given in Table 12.4. A comparison of the per cent giving the favorable responses (especially the most favorable response—marked with an asterisk) for items 1–4 indicates that the employees in plant B felt much more on the same "wave length" with their supervisors than was true of employees in plant A. The responses to item 5 indicate

---

[11] Habbe, *op. cit.*

TABLE 12.4

Responses of Employees in Two Plants to Questions on Attitude Questionnaire

Question and response	*Per cent giving response*	
	Plant A	Plant B
1. In your own experience, who has been of most help in keeping you up to date on company matters?		
My steward	16	22
My foreman	37*	47*
A fellow worker	32	21
One of the top men	15	10
2. When you are asked to do something by your foreman, are you told why?		
Almost never told	24	24
About half the time	38	21
Almost always told	38*	55*
3. Can you talk things over with your foreman when you want to?		
Yes, I always can	34*	56*
Usually I can	42	27
He is generally too busy	6	4
Usually he doesn't want to be bothered	13	13
I can hardly ever talk with him	5	0
4. Have you been able to get your ideas up to the top men?		
Almost always	21*	49*
Sometimes	31	28
Hardly ever	32	13
Not interested in doing so	16	10
5. Do you feel a part of your company?		
I feel I really belong	29*	62*
I feel I just work here	42	14
Sometimes I feel one way, sometimes the other	29	24

* This is the most favorable response. A comparison for the two plants on this response points up the differences between them.

Source: Habbe, *op. cit.*, pp. 36–38.

a much greater feeling of "belongingness" on the part of employees in plant B as contrasted with plant A (62 per cent *vs.* 29 per cent). A generally more favorable attitude was expressed by employees of plant B in answering another question (not shown in the table) about their general feelings toward the company. Forty-five per cent of the employees in plant B said it was "one of the best," whereas in plant A only 20 per cent gave this response.

It should be pointed out here, as in certain previous sections, that one needs to be careful in drawing positive cause-and-effect conclusions from data such as these. The differences in attitudes between these plants may not be due to the fact that one had a "communication program" and the other did not. It may be that the general frame of reference of the management of plant B contributed to the differences. Perhaps the "communication program" was only one aspect of a broader policy involving employee relations. If this were the case, however, one would be strongly inclined toward the conviction that employee communication plays an important part in the management of personnel.

Another study dealing with an aspect of communications was reported by Mann and Baumgartel[12] in their analysis of absences in an electric power company. Responses to questions in an attitude questionnaire were analyzed in relationship to absenteeism. The results for two groups of men (white-collar and blue-collar) are given in Fig. 12.3. The absenteeism data are average times absent for men within work groups. The bars represent the per cent within work-groups of the specified absenteeism indexes who gave the response indicated. These results indicate that supervisors who create an atmosphere that makes for free and easy discussion of work problems or personal problems tend to have fewer absences in their work groups than do those who do not create such an atmosphere.

Nature of Communicative Ability. The survey of corporation presidents mentioned above included certain questions regarding the nature of communicative ability, with particular reference to whether it is a "talent" or whether it can be learned. Of the 51 presidents, 46 expressed the opinion that communicative ability is a combination of talent and skill, and 35 expressed the opinion that it can be learned. Their conviction that such skills can be acquired

---

[12] F. Mann and H. Baumgartel, *Absences and Employee Attitudes in an Electric Power Company*, Survey Research Center, University of Michigan, December 1952.

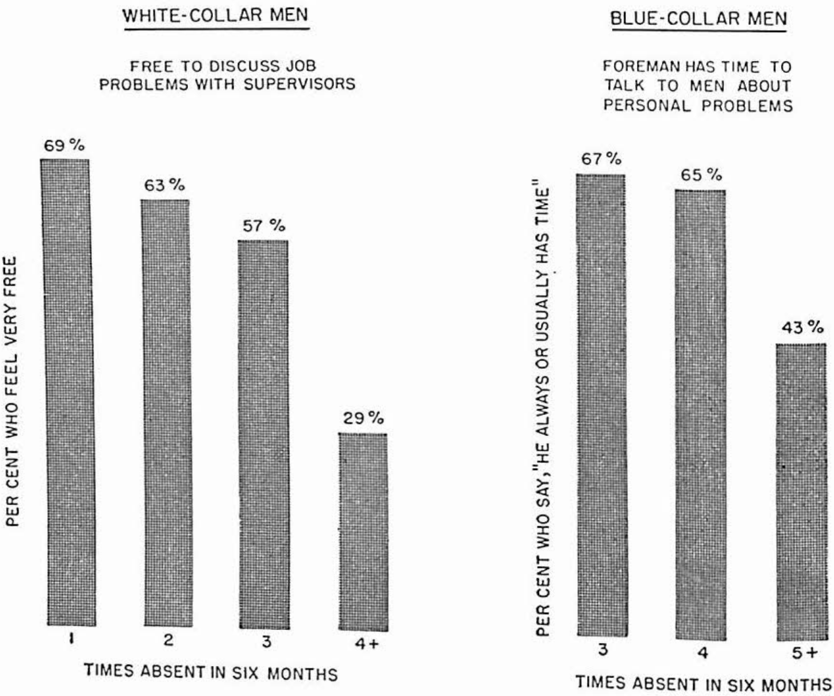


FIG. 12.3. Relationship between absenteeism and freedom of supervisor-employee communications for white-collar and blue-collar men in an electric power company. (From Mann and Baumgartel, op. cit.)

in part through training is indicated by the fact that 49 of them said they believed training in communication skills for management personnel was desirable (item 7 of Table 12.2). More of them felt such training was more necessary for middle management and first-line supervisors (44 and 45 respectively) than for top management (33), as indicated in item 8.

Readability of Written Communications. The number of written communications in business and industry, if it were known, would be astronomical. For such pieces of paper to "communicate" their messages, however, they must be written in a manner that is understandable to those who read them. A method for measuring readability of written material has been developed by Flesch.[13]

---

[13] R. Flesch, "A New Readability Yardstick," *Journal of Applied Psychology,* 1948, **32,** 221–233.

In the measurement of reading ease, Flesch takes into account the following elements: (1) average sentence length in words, (2) average word length in syllables, (3) average percentage of "personal words" (such as personal pronouns), and (4) average percentage of "personal sentences" (such as spoken sentences, questions, commands, requests). These elements typically are determined for a sample of the total written material and are used for deriving two indexes, as follows:

RE (reading ease) $= 206.835 - .846\ wl$ (word length) $- 1.015\ sl$
(syllable length)
HI (human interest) $= 3.635\ pw$ (personal words) $+ .314\ ps$
(personal sentences)

These indexes range from 0 to 100 for almost all samples of writing analyzed. In general, the "reading ease" and "human interest" scores tend to follow the patterns shown in Tables 12.5 and 12.6.

TABLE 12.5

Pattern of "Reading Ease" Scores

"Reading Ease" Score	*Description of Style*	*Typical Magazine*	*Syllables per 100 Words*	*Average Sentence Length in Words*
0 to 30	Very difficult	Scientific	192 or more	29 or more
30 to 50	Difficult	Academic	167	25
50 to 60	Fairly difficult	Quality	155	21
60 to 70	Standard	Digests	147	17
70 to 80	Fairly easy	Slick-fiction	139	14
80 to 90	Easy	Pulp-fiction	131	11
90 to 100	Very easy	Comics	123 or less	8 or less

Source: Flesch, *op. cit.*, Table 5.

TABLE 12.6

Pattern of "Human Interest" Scores

"Human Interest" Score	*Description of Style*	*Typical Magazine*	*Percentage of Personal Words*	*Percentage of Personal Sentences*
0 to 10	Dull	Scientific	2 or less	0
10 to 20	Mildly interesting	Trade	4	5
20 to 40	Interesting	Digests	7	15
40 to 60	Highly interesting	*New Yorker*	11	32
60 to 100	Dramatic	Fiction	17 or more	58 or more

Source: Flesch, *op. cit.*, Table 6.

An example of the reading difficulty of a section in a typical management-union agreement with a simplified version of the same section written in a much more readable manner has been published by Lauer and Paterson.[14] The following passage is reproduced exactly as it appeared in the original contract:

> The Employer shall not discharge any employee without just cause and shall give at least one warning notice of the complaint in writing to the employee and to the Union, except that no warning notice need be given to an employee before he is discharged if the cause of such discharge is dishonesty, major violation of company rules that do not conflict with this agreement, or drinking while on duty.

The contract from which this passage was taken has a Flesch Reading Ease score of 28.3—which classifies it as "Very Difficult" and, according to Flesch, requires a college education or its equivalent for adequate understanding.

Here is a simplified version of the foregoing contract provision:

> The Employer cannot fire an employee without just cause. He has to write the employee and the Union and warn them at least once about the fault. No warning has to be given to an employee before he is fired for these causes:
>
> (*a*) He is dishonest.
> (*b*) He commits a major violation of company rules which do not conflict with this agreement.
> (*c*) He drinks while on duty.

The simplified version of the contract, from which the above passage was taken, has a Reading Ease score of 72.7, which classifies it as "Fairly Easy." According to Flesch, people with only a sixth grade education can readily read and understand this version. The simplification has been accomplished by breaking up the original, very long sentence into short sentences. In other words, what the original passage says in 70 words (all in one sentence), is said more effectively in the simplified version in only 66 words (broken into six short sentences).

In a practical situation these indexes can be used to evaluate the reading ease and human interest of written communications. If they are then compared with the reading abilities of those who are to read the communications it is possible to see whether the communications are within, or above, the reading abilities of the people concerned. There seems to be evidence that many of the written communications in industry *are* above the reading-ability levels of employees, as indicated by Colby and Tiffin,[15] Paterson and Jenkins,[16] and Raney.[17]

---

[14] Jeanne Lauer and D. G. Paterson, "Readability of Union Contracts," *Personnel,* 1951, 28, 36–40.

Various surveys have been made of the readability of different kinds of written communications in industry. In a study of 50 house organs (employee magazines), for example, England[18] reports that the average reading-ease score was 48.4. This means that the material was "difficult" to read and understand. In particular it was written at the level found in textbooks and semiscientific journals, and could be read easily only by high school graduates. Since the majority of adults do not have that level of education, it is apparent that such house organs would be "over the heads" of many employees. England offers the following suggestions for the more suitable writing of house organs:

1. Define your purpose for writing; what do you want the story to tell?
2. Present your ideas in logical order; don't jump from one idea to another.
3. Use words of one or two syllables whenever possible.
4. Keep the vocabulary familiar; use Thorndike's dictionary when in doubt.
5. Use simple sentence structure.

Another common type of communication is the union-management agreement. An analysis of the readability of 59 union-management agreements was made by Tiffin and Walsh.[19] Fig. 12.4 shows the distribution of reading-ease scores of these agreements, along with a comparison of the estimated per cent of adult readers having attained the indicated educational level. This indicates very clearly that 96.6 per cent of the agreements were written at levels that required high school graduation or some college education for easy

---

[15] A. N. Colby and J. Tiffin, "The Reading Ability of Industrial Supervisors," *Personnel*, 1950, **27**, 156–159.

[16] D. G. Paterson and J. J. Jenkins, "Communications Between Management and Workers," *Journal of Applied Psychology*, 1948, **32**, 71–80.

[17] E. T. Raney, "How Readable Are Your Employee Publications?" *Personnel Psychology*, 1949, **2**, 437–459.

[18] A. O. England, "Employee Magazines Are Hard to Read," *Personnel Journal*, 1951, **30**, 94–97.

[19] J. Tiffin and F. X. Walsh, "Readability of Union-Management Agreements," *Personnel Psychology*, 1951, **4**, 327–337.

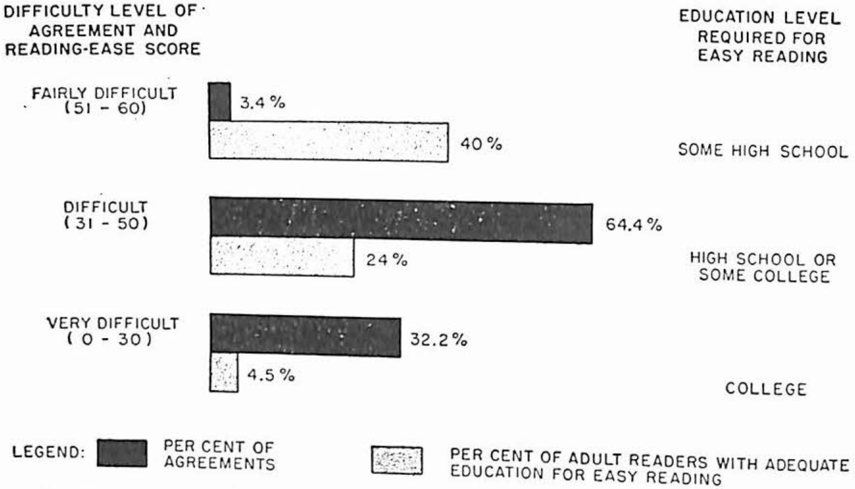


FIG. 12.4. Relationship of reading ease of 59 union-management agreements and per cent of adult readers having educational levels required for easy reading. (Adapted from Tiffin and Walsh, op. cit.)

understanding, whereas only about 28.5 per cent of the potential adult readers would be expected to have such education.

It might be added, incidentally, that the agreements with large unions were more difficult to read than those with smaller unions. The study, however, did not reveal any reasons for this.

It seems abundantly clear that communications comprise an integral link in the chain of human relations in industry. The success of a communications program, however, may in part depend upon the extent to which the communications are expressed in such a way that they are easily understood by those to whom they are addressed.

EMPLOYEE PARTICIPATION

In recent years much has been said—pro and con—about the participation of employees in some of the decisions relating to the work situation. Viteles,[20] for example, summarizes the conclusions reported from some of the research on employee participation, as follows: ". . . major outcome of such research is the conclusion that *employee participation in decision-making* in a *democratic atmosphere* created by 'permissive' leadership, facilitates the development of 'internalized' motivation, and serves to raise the levels of the employee production and morale."

---

[20] M. S. Viteles, *Motivation and Morale in Industry* (New York, W. W. Norton & Co., 1953), p. 164.

On the other hand, there are those who on various grounds question the desirability of participation. Dale,[21] for example, in reporting the results of a survey among management representatives (as well as union officials and employees) regarding their experiences in formal labor-management cooperation efforts, indicates that the complaints by management about such collaboration were made on the following three grounds: (1) labor was not acting in "good faith"; (2) management functions were challenged or invaded; and (3) the economic costs of cooperation outweighed the gains. It might be added that of the 201 companies with formal cooperation plans that were covered by the survey, the managements of about two-thirds indicated that they were "pleased" with the results of cooperation, with the other one-third indicating varying degrees of neutrality, disappointment, or a "wait and see" attitude with respect to the results.

Let us reserve any evaluation of these conflicting views, however, until we have discussed more what we *mean* by employee participation, and until we have discussed certain specific cases.

Types of Employee Participation. There are two general types of employee "participation" in industry—namely, what might be considered to be *formal* and *informal.*

Formal participation consists of some plan for labor-management cooperation that is, to some degree, recognized as a *modus operandi* between management and employees, frequently through a union. The degree and character of such "cooperation" varies, as does the field of cooperation. In the survey of 201 plans reported by Dale,[22] for example, there were 22 areas of cooperation, as shown in Table 12.7. Among the most commonly mentioned areas of such plans were accident prevention, elimination of waste and defective work, furthering labor understanding of policies, attendance, employee insurance plans, quality control, and job evaluation.

The informal type of employee participation occurs more typically at the work-group level, where the supervisor develops the opportunity for the group to make certain types of decisions affecting the group as a whole. Usually the matters on which such group decisions are made are those which are completely within the prerogatives of the supervisor, and are not company-policy matters. It needs to be reiterated, however, that the distinction between formal and informal participation is one of degree.

---

[21] E. Dale, *Greater Productivity Through Labor-Management Cooperation,* Research Report No. 14, American Management Association, 1949, p. 26.

[22] E. Dale, *op. cit.*

In characterizing group decision-making, Maier[23] presents the the following, indicating both what it *is,* and what it is *not.*

Group Decision Is Not	*Group Decision Is*
1. Abandoning control of the situation.	A way of controlling through leadership rather than force.
2. A disregard of discipline.	A way of group discipline through social pressure.
3. A way of giving each individual what he wants.	A way of being fair to the job and all members of a group.
4. A way of manipulating people.	A way of reconciling conflicting attitudes.
5. A way of selling the supervisor's ideas to a group.	Permitting the group to jell on the idea it thinks will best solve a problem.
6. Sugar-coated autocracy.	A way of letting facts and feelings operate.
7. A matter of collecting votes.	Pooled thinking.
8. Consultative supervision in which mere advice is sought.	Cooperative problem-solving.
9. A way of turning the company over to employees.	A way of giving each person a chance to participate in things that concern him in his work situation.
10. Something anyone can do if he wishes.	A method that requires skill and a respect for other people.

Examples of Employee Participation. One of the most widely known studies relating to employee participation was carried out at a pajama manufacturing company, the Harwood Manufacturing Corporation, Marion, Virginia, as reported by Marrow[24] and by Coch

---

[23] N. R. F. Maier, *Principles of Human Relations* (New York, John Wiley and Sons, Inc., 1952).

[24] A. Marrow, "Industrial Psychology Pays in This Plant," *Modern Industry,* July 15, 1948, **16,** No. 1, 67 ff.

TABLE 12.7

Areas of Labor-Management Cooperation Plans in 201 Companies

Area	*Number of times mentioned*
Accident prevention	163
Elimination of waste and defective work	147
Furthering labor understanding of policies	145
Regular attendance	115
Employee insurance plans	110
Quality control	105
Job evaluation	102
Physical working conditions	101
Lateness	100
Maintenance of tools, fixtures, etc.	98
Employee health	97
Methods improvement	96
Discipline control	88
Training, apprenticeship, and induction programs	85
Labor turnover	69
Setting output standards	68
Incentive systems	67
Production planning	59
Promotional programs	53
Utilization of machinery	51
Employment stabilization	49
Technological changes	48

Source: Adapted from Dale, *op. cit.*, p. 24.

and French.[25] Marrow was the president of the corporation, and himself a psychologist. Because of changes in style trends it became necessary to change the design of the garments being made. A change in garment design, incidentally, necessitates changes in work assignments and the relearning of the new assignments, usually accompanied by reduced piece-rate earnings until the operators get back up to their pre-change levels. Sometimes they never do get back up to their early levels. Because of these factors, there typically is considerable employee resistance to changes in designs.

Since circumstances made job changes necessary, an opportunity existed for an experiment in group participation. In this experiment, four groups of sewing-machine operators were formed, all four groups being matched for the difficulty of the new jobs, for the amount of change in their jobs, and for the level of productivity before the experiment. Three variations in the manner of dealing with these groups were used in the experiment, as follows:

---

[25] L. Coch and J. R. P. French, Jr., "Overcoming Resistance to Change," *Human Relations*, 1948, I, 512–532.

Control group. This group was changed to the new jobs by the normal factory procedure. They were given an explanation of why a change in job methods was necessary, what the new job would be like, and what the new piece rates would be.

Experimental group 1. A moderate degree of democratic participation was used with this group. They received more information and explanation about the need for change, and they were provided the opportunity to choose representatives who, in turn, participated in designing the new job, setting the new piece rate, and later in training the remaining members of the group. The feeling of participation on the part of the group was such that they spoke of the new job and new piece rates as "our job" and "our piece rates."

Experimental groups 2 and 3. These groups participated directly in designing the job and setting the new piece rates, these having a greater degree of participation than did experimental group 1, which "participated" through chosen representatives.

Records of performance of these groups are shown in Fig. 12.5. The "units per hour" is an index based on standard units of work determined by time and motion study methods; thus, it provides a reasonably equivalent measure of output regardless of differences among jobs. The results, as presented in Fig. 12.5, were quite dramatic. It can be seen for the period of time before the change that all groups were approximately equal. After the change the control group dropped to an index of about 50—the usual occurrence after such a change. Many of those in this group quit their jobs. Group 1, which had participation through chosen representatives, dropped to begin with, but consistently recovered over a period of time, ultimately achieving an index of about 65. Groups 2 and 3, which had direct participation, in turn, experienced practically no drop, and continued to increase their performance until they finally levelled off at an index in the low 70's. Not only was the change in productivity after change proportional to the degree of participation of the three groups, but the turnover rates and amounts of aggression expressed against management were inversely proportional to the degree of participation.

A summary of data for several companies with formal labor-

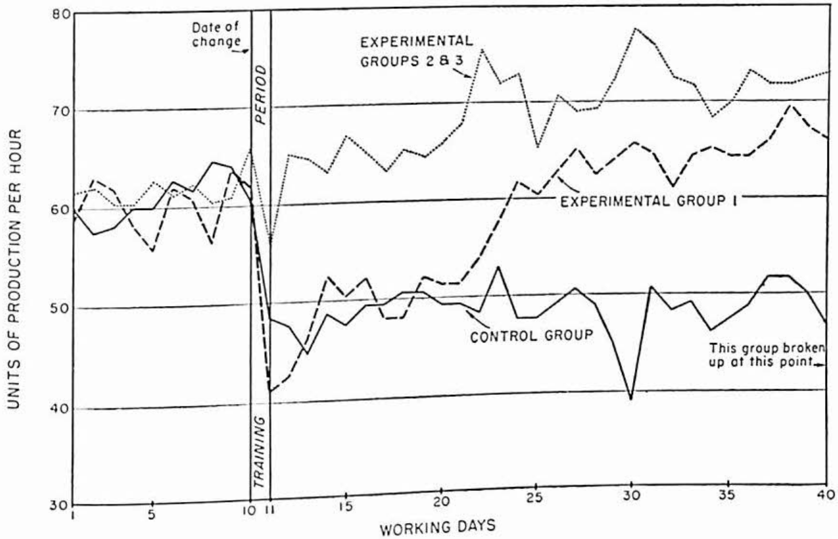


FIG. 12.5. Relationship between degree of participation and work performance of four groups of sewing machine operators. See text for description of participation of groups. (Adapted from Marrow, op. cit.)

management cooperation plans was prepared by Dale[26] from War Production Board records. This summary included absenteeism rates for the several companies before and after labor participation; the rates are reproduced in Table 12.8. These data show reductions in absenteeism in all nine companies for which data are reported.

In connection with situations in which more informal participation has been used, Maier[27] cites the case of a telephone repair crew consisting of 12 men. The number of repair "visits" varies with the types of repairs to be performed, but these differences tend to average out over a period of time. The average for the crew in question for a six-month period had been between 8 and 9 per day, as compared with the company average of 10.8 visits per day.

In part because of their below-average performance, the foreman held a group meeting to ask his men whether they would care to discuss, as a group, any ideas on how their work could be better coordinated, and any obstacles and difficulties that they thought remains to be learned about the variables that are related to employee behavior.

---

[26] Dale, *op. cit.*, p. 56.

[27] N. R. F. Maier, *Principles of Human Relations, op. cit.*, pp. 225–228.

Granting, however, that much remains to be learned about employee participation, it is obviously not possible to stop the wheels of industry to wait until some further answers are forthcoming from research. In attempting to assess, on presently available evidence, the appropriateness of employee participation as a human relations method, there seems to be more on the positive side of the ledger than on the negative side. As with many other personnel practices, participation needs to be used judiciously. In particular there seem to be at least two types of restrictions on its potential use. One of these relates to the work group itself. The nature and composition of the group, its history of working relationships, its functions, and so on will temper the potential applicability of group participation in decision-making. The second restriction is that the area of decision-making should have a bearing on the appropriateness of group methods of making decisions. There are, for example, certain functions that almost by their nature are the prerogatives and even the inescapable obligations of management. To avoid such obligations as the selection of men for promotion to most management positions, for example, amounts almost to the abrogation of one's responsibilities. On the other hand, there are certain functions or situations, such as some of those mentioned in the cases cited above, where the decisions can more appropriately be the consequence of group processes.

GROUP UNITY

The cohesiveness of a work group can have a powerful influence on the behavior of the individuals within the group. However, as pointed out by Wilensky[30] and Viteles,[31] as well as others, this influence can operate in either of two directions.

Differences in Group Behavior. The bleak, unsavory side of the picture was revealed by certain of the studies in the Hawthorne plant of the Western Electric Company, as reported by Roethlisberger and Dickson,[32] in which they point up the influence of the work group on the behavior of individuals in restriction of output. This was particularly the case among the employees in the bank wiring observation room, which was one of the groups studied. It was the prevailing opinion of those in this shop that a person who turned out too much work was a "ratebuster."

---

[30] Wilensky, *op. cit.*

[31] Viteles, *op. cit.*

[32] F. J. Roethlisberger and W. J. Dickson, *Management and the Worker* (Harvard University Press, 1939).

Drucker,[33] in discussing output restrictions, expresses the opinion that open restrictions on output provided openly under the provisions of restrictive union rules and "featherbedding" are "only the part of the iceberg that is above the water. Much more important are the invisible, unwritten, informal restrictions decreed by the custom and common law of every plant," as reflected in the tacit setting of production quotas (by the work group) which a worker would be "ill-advised" to exceed.

The more positive effects of work-group unity are revealed by various studies in which it has been demonstrated that there was a relationship between feelings of group cohesiveness and increases in work performance, or other desirable changes in employee behavior, such as reduction of absenteeism. In the study of employees in an insurance company as reported by Katz, MacCoby, and Morse,[34] for example, the employees in the high-producing sections were more strongly "identified" with their groups than were employees in the low-producing sections. There was also evidence from the electric power company investigation reported by Mann and Baumgartel,[35] of a distinct relationship between group unity and absenteeism. Some evidence of this type is shown in Fig. 12.7. In particular this shows the per cent of employees who expressed favorable attitudes toward their work groups, in relationship to the absenteeism averages of their groups. The patterns of relationship are somewhat similar for both blue-collar and white-collar men; in work groups with high absence rates there were fewer who expressed feelings of "group unity" than in work groups with low absence rates.

Since feelings of group unity presumably can influence behavior in either a *desirable* or *undesirable* direction (in terms of employee performance), it is important to try to understand *why* such differences occur. The most rational explanation seems to be in terms of group *goals*. When the group has developed goals that are oriented

---

[33] P. F. Drucker, *The New Society* (Harper & Brothers, 1950), p. 83.

[34] Katz, MacCoby, and Morse, *op. cit.*

[35] Mann and Baumgartel, *op. cit.*

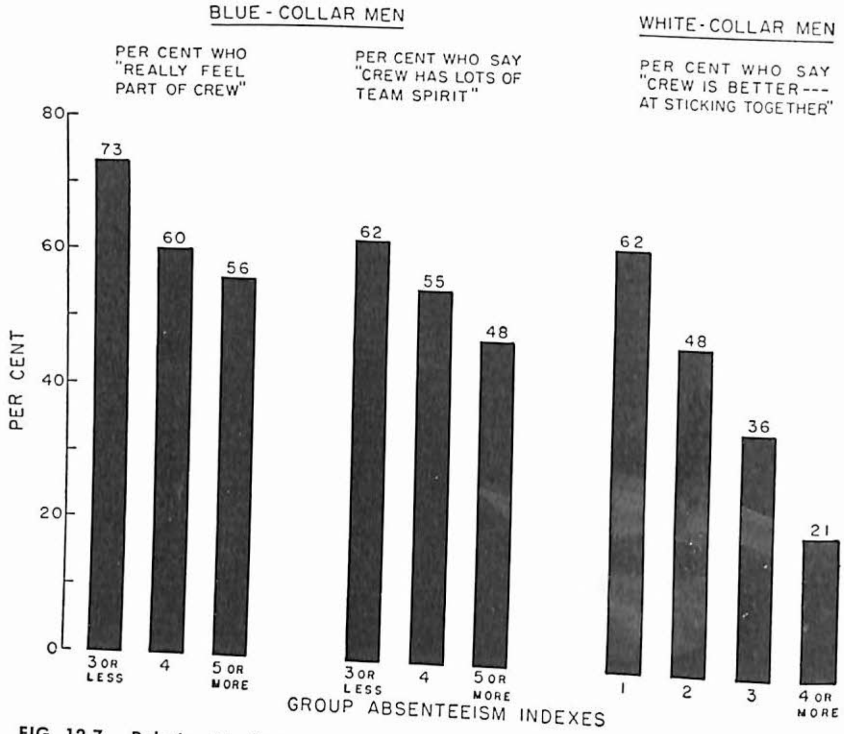


FIG. 12.7. Relationship between absence indexes and attitudes of employees toward their work groups in electric power company. (Adapted from Mann and Baumgartel, op. cit.)

primarily in terms of their own self-interest, the motivation to achieve those objectives might be in a direction that is at odds with the objectives of the organization. This would be the case, for example, if a work group were seeking to prevent or delay layoffs by stretching the work out, or to prevent the company from changing the "rate" on a job by intentionally restricting output. On the other hand, when the goals of the group correspond with those of the organization, the behavior is more likely to be in the direction of the organizational objectives.

The Development of Group Unity. Although there may be those who would argue with this point of view, there seems to be adequate justification for proposing that one objective of a human relations program is that of developing a situation in which the employees have a strong sense of group unity, and in which they have goals in common with those of the management. In fact, Worthy[36] expresses the opinion that "One of the measures of the effectiveness of an organization is the extent to which the aims of the informal organization correspond with those of the formal." Continuing in this vein he makes the following statement:

> The manager's problem is not that of 'bending the informal work group to the purposes of the enterprise.' Rather, it is so conducting the enterprise that the relationship between its purposes and those of the primary work group are clearly apparent.

Such unity of purpose cannot be accomplished by using a ready-made formula. Rather, it is more typically the end-result of many factors, including those that have been mentioned above, such as the nature of the supervision received, the freedom for communications, and the opportunity for some form of participation in decision-making in certain areas.

Formation of congenial work groups. Although much can be done by a supervisor in building a group spirit with an existing work group, there sometimes is the opportunity to form work groups of individuals who indicate that they would like to work together.

A case of this type is reported by Van Zelst[37] in connection with a construction job in which rows of identical houses were being built. The 38 carpenters and 36 bricklayers on the job had been working together for at least five months, and so they were well acquainted with each other's personality and skill.

Sometime after the construction job was started, an experiment was begun, in which the workers were regrouped into groups of two (or, in some cases, four) on the basis of personal choices of work partners. In doing this, each person was asked to list his first, second, and third choices for a partner. All but eight were paired with one of these three, and these eight "isolates" were paired until they could be successfully incorporated into a mutual-choice group.

Objective records were available on labor cost, materials cost, and turnover for the 9-month period *before* this experimental period, and for the 11-month experimental period itself. These data are tion operations, they have their limitations, too. This is evident in the fact that the performance of some inspectors is below desired standards. Further evidence to support this statement will be given later.

---

[36] J. C. Worthy, "Comments on Mr. Wilensky's Chapter," in *Research in Industrial Human Relations, op. cit.*, pp. 52–53.

[37] R. H. Van Zelst, "Sociometrically Selected Work Teams Increase Production," *Personnel Psychology*, 1952, **5**, 175–185.

Although human error can occur in inspection operations, it is frequently possible to reduce this error, thereby increasing the effectiveness of the inspection process. In fact, the major purpose of this chapter is to demonstrate, by example, how systematic study of inspection jobs can contribute to the solution of certain human problems involved with those jobs. It is recognized that the specific conclusions reached in the cited examples may not apply directly to inspection of other products. The basic methods described, however, might well apply to corresponding problems with other inspection operations. It must be emphasized that the present discussion deals with *methods*, especially as related to problems of *placement* of inspectors, of the *methods* of inspection, and of the *training* of inspectors.

MEASURING THE ACCURACY OF INSPECTORS

In carrying out a study for any of the above purposes it usually is necessary to have some measure of the job performance of employees. Such a measure is used as a *criterion*. In many production jobs it is possible to determine the relative performance of employees within a group from records of production or earnings, or from supervisory ratings. In many inspection jobs, however, records do not reflect the relative performance of inspectors, and ratings in some cases have been found to be related more to *quantity* of materials inspected than to the *ability* of the inspectors to identify products that should be rejected.

Thus, in studying some inspection jobs a special problem exists in obtaining information about the relative abilities of inspectors. One possible approach to this is to use a *job sample test*. A few examples of such tests will be described.

Accuracy in Tin-Plate Inspection.[1] One of these examples concerns the job of tin-plate inspector. Since this job will be referred to later, it will be described briefly.

---

[1] This study of tin-plate inspection is based on an article published by J. Tiffin and H. B. Rogers, "The Selection and Training of Inspectors," *Personnel*, 1941, 18, 14–31.

The job of tin-plate inspecting, ordinarily called *assorting*, is essentially an inspection for appearance that is made while the inspector turns the sheets of tin plate from one stack to another. As the sheet is turned, the inspector makes a decision from the appearance and feel of the sheet whether it is a prime or a second or contains one of a number of possible defects. The standard size sheet is approximately 28 by 30 inches, though this varies with the specific order. In the plant studied, the work is done by women assorters. Supervisors have generally felt that inspectors on this job do not reach their maximum performance until they have had approximately six months of experience. The inspectors are paid on a straight hourly rate.

The mechanics of wrapping and shipping made it impractical to reinspect samples of tin-plate in order to evaluate the performance of the inspectors. Under the circumstances, a job sample test was used. This job sample was developed after considerable preliminary investigation. In its final form it consisted of a coded stack of 150 sheets. This coded stack was made up of 61 prime sheets (sheets satisfactory in every respect) and second sheets (sheets containing a minor surface blemish or uneven coating of tin), 30 sheets containing appearance defect No. 1, 26 sheets containing appearance defect No. 2, 13 sheets containing appearance defect No. 3, and 20 sheets containing a weight defect. The prime and second sheets were included in a single category for the purpose of the experiment because with repeated assortings the prime sheets tend to become scratched and thus to become seconds. The other defective plates included in the coded stack remained the same regardless of the number of times the stack was assorted.

The defects will be referred to by number rather than by name because it is the method of measuring the accuracy of the inspectors rather than their accuracy on specific defects that is of primary importance.

Each sheet included in the coded stack was carefully selected in advance and was known definitely to be either a prime or a second, or to contain one of the defects selected for study. The 150 sheets were numbered in random sequence. As the inspectors assorted the sheets, they called aloud their judgment of each sheet, whether it was a prime or a second or contained one of the defects, and, if so, which one.

As the inspectors assorted the coded stack, they were allowed to set their own speed or rate of inspection. This procedure resulted in very great differences in speed. The time required for the various girls varied from eight to forty-eight minutes. It was first thought that this uncontrolled time factor would be a marked handicap to this method of measuring inspector accuracy. After careful consideration, however, it was thought best to retain an indefinite-time-limit procedure because only under such circumstances would it be possible to determine for each girl the maximum accuracy of which she was capable. In other words, the coded stack test was deliberately given under conditions that would cause every girl to bring to bear every bit of inspecting knowledge she had and that therefore would allow her to reach a level of accuracy that was close to, if not equal to, the maximum accuracy of which she was capable.

As each operator assorted the 150 sheets her judgments were recorded by an observer. She was also timed with a stop watch. This procedure resulted in data from which it was possible to determine both the speed and the accuracy of the assorters studied. A total of 150 inspectors took this job sample test.

The coded stack was scored for each girl according to the total accuracy on the 150 sheets and also for the accuracy on the specific defects. This procedure resulted in five specific measurements of accuracy for each girl. It also resulted in a measurement of the reliability of each of the five methods of scoring the coded stack test. The five resulting measurements with their respective reliabilities are summarized in Table 13.1.

TABLE 13.1

The Five Measurements of Accuracy Yielded by the Coded Stack, with the Reliability of Each

Method of Scoring	*Reliability*
Mixed sheets	+.90
Appearance defect No. 1	+.86
Appearance defect No. 2	+.87
Appearance defect No. 3	+.68
Weight defect	+.74

The reliability figures given in Table 13.1 were obtained by correlating correct inspection on odd *vs.* even items for each defect, as described on p. 93. The reliability indicates the extent to which repeated or duplicate measurements of each girl by means of the coded stack test would result in the same score for her for the defects in question. It may be seen from Table 13.1 that the reliabilities vary from .68 to .90. Although these reliabilities are not as high as might be desired, they compare favorably with the reliabilities of many other industrial criteria that have been used to measure the successfulness of employees.

The results of the coded stack test as administered to the 150 girls are summarized in graphic form in Fig. 13.1. The six curves shown in this figure are frequency distributions obtained from the 150 inspectors.

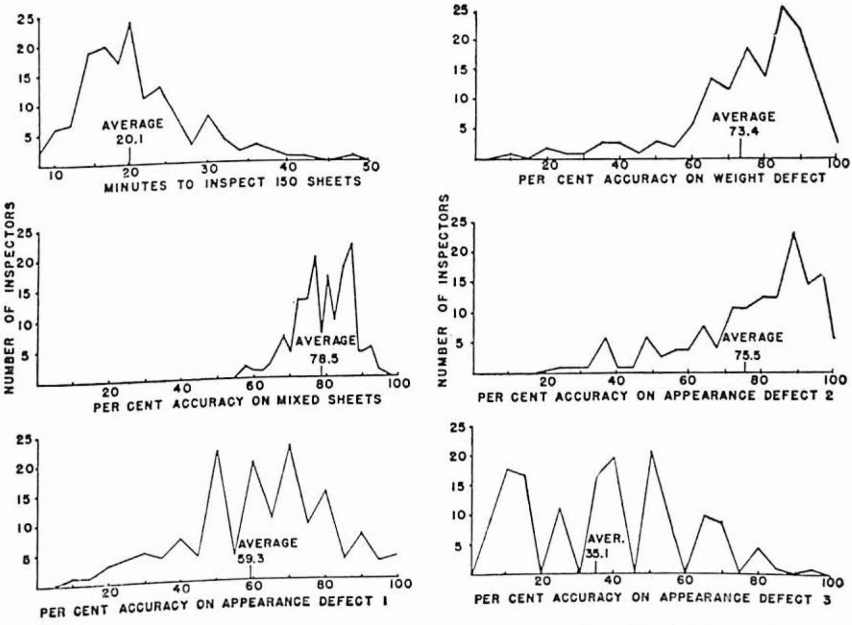


FIG. 13.1. Distributions of results on a coded stack test of 150 tin-plate inspectors.

The time curve in the upper left-hand corner indicates the number of girls completing the 150 sheets of tin plate in each of the times specified along the base line. The most rapid assorter, it will be noted, completed the test in eight minutes, whereas the slowest required forty-eight minutes. The curve is highest over the region of fourteen to twenty-one minutes, indicating that a plurality of the girls required between fourteen and twenty-one minutes to assort the 150 sheets under the conditions of the test. The average time required, as indicated in Fig. 13.1, is 20.1 minutes.

In the curve immediately below this, the ordinate again represents the number of girls and the base line represents the percentage of accuracy for the entire stack of 150 sheets. The average accuracy for the 150 sheets under the conditions of this test was, as shown in Fig. 13.1, 78.5 per cent. This does not mean that the accuracy of these assorters in a normal assorting situation is only 78.5 per cent. The assorting skill required to inspect the coded stack is greater than that required in a normal assorting situation, because more defective plates of various types are included in the test than are found in the run of the mill and the defective plates were not included in any sequence or groupings. The test as given measures differences among the girls in their skill on this job but probably gives to each girl a somewhat lower score—that is, a poorer accuracy—than she customarily reaches in routine daily assorting.

The remaining four curves in Fig. 13.1 give similar distributions for the four specific defects studied. Fig. 13.1 shows, among other things, that for the four defects studied the average inspector is most accurate on appearance defect No. 2 (75.5 per cent), and least accurate on appearance defect No. 3 (35.1 per cent). The data also show a great variation among the various girls in their accuracy on any specific type of defect. For example, 8 girls out of the 150 detected only 5 per cent of appearance defect No. 3, whereas one girl spotted 95 per cent of the sheets containing this defect. There is nearly as wide a variation in accuracy for the other four defects studied.

It will be noticed that the average accuracy on mixed sheets is higher than the accuracy on any specific defect. This difference is due to the fact that a majority of the sheets included in the coded stack were primes or seconds and that the acceptable sheets were seldom called defective, whereas the defective sheets were frequently called primes or seconds. In other words, there was a tendency to overlook the defects. The accuracy for the total stack, therefore, which included a majority of primes and seconds and a minority of defective sheets, was found to be higher than the accuracy on any specific defect studied.

The next question investigated was the relation between speed and accuracy in detecting the different types of defect. The results

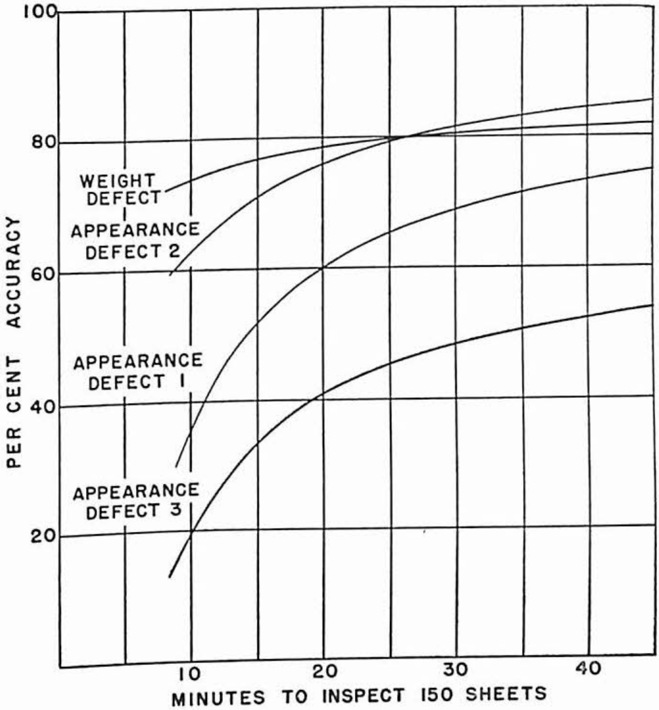


FIG. 13.2. Relationships between speed of inspection and accuracy in detecting four defects.

for this part of the investigation are summarized in graphic form by the curves in Fig. 13.2.

It is interesting to note that in detecting off-weight sheets the accuracy is about the same whether the test is completed in ten minutes or in forty minutes. In other words, accuracy in detecting off-weight sheets for the 150 girls tested is not appreciably affected by speed of inspecting within the time limits found for these girls. To some extent the same general situation is true for appearance defect 2, although in detecting this defect we note a beginning tendency for accuracy to decrease with increased speed. This tendency is still more pronounced in the case of appearance defects 1 and 3. For example, we find that when the stack is assorted in ten minutes, only 20 per cent of the sheets containing appearance defect No. 3 are spotted and only 35 per cent of those containing appearance defect No. 1.

On many inspection jobs certain defects may be spotted equally well at either fast or slow rates of inspection, but other defects cannot be spotted with satisfactory accuracy above a certain inspection speed. These curves clearly reveal this situation for this inspection job. Defects of weight are noted almost as well when the 150 pieces are inspected in ten minutes as they are when forty minutes are used to inspect this number of sheets. It is important for the head of an inspection department to know how accuracy is related to varying speed. When he possesses this information he is better able to influence the rate of work of the inspectors so as to achieve satisfactory quality without wasting time in over-careful inspection.

The coded stack used in this study furnished not only a means of measuring the accuracy of the inspectors but also data from which the relationships between speed and accuracy plotted in Fig. 13.2 were determined.

Accuracy in the Use of Precision Instruments. Other examples of job sample tests with inspectors include a study by Lawshe and Tiffin.[2] In particular, this investigation dealt with the measurement of metal parts with precision instruments. Many precision instruments used in machine shops require that the operator judge the "feel," "tension," "drag," or other characteristics; yet, in spite of all that is known about the variability of human judgments, little attention has been given to the importance of such variability as it may affect the accuracy of measurements.

One of these examples is from a plant engaged in the manufacture of variable-pitch propellers for aircraft. Approximately 200 people were employed in the inspection department. Each of the various classifications of inspection jobs was analyzed to determine what precision instruments were used on the job, and what tolerances were demanded. Frequency counts were then made to determine which instruments or combination of instruments were used in the largest number of classifications and by the largest number of employees. On the basis of this count, twenty instruments and combinations were chosen as being most important in this particular plant.

A room was set aside as a dimensional control laboratory and twenty booths or inspection stations were set up. Each booth was numbered, and in each were placed one of the twenty instruments, a standard part from the plant, and a simplified working drawing that indicated one dimension to be measured with the instrument provided. When an employee entered the room, the attendant determined his job classification and provided him with an appropriate work-sheet for each of the stations containing work samples from his job. Each employee was tested on only those instruments that he used on his particular job. He was encouraged to make five measurements and then to record his best judgment concerning the dimension. The readings thus obtained were compared with so-called "true" dimensions, which were determined by means of ultra-precision instruments in combination with Johanssen blocks. Instruments utilized in the performance testing were checked and adjusted periodically to insure constancy.

---

[2] C. H. Lawshe and J. Tiffin, "The Accuracy of Precision Instrument Measurement in Industrial Inspection," *Journal of Applied Psychology*, 1945, **29**, 413–419.

Results obtained at eleven of the twenty stations are presented in Fig. 13.3. The particular stations selected for illustration were chosen in terms of general familiarity with the instruments used and not because of any peculiarity in the findings; they are truly representative.

In Fig. 13.3, the open bars represent the percentage of inspectors tested who obtained readings within the established tolerance. The solid bars represent the percentage of persons tested who failed to meet the standard. As already stated, not all the inspectors were tested at each station; instead, the sample contains only those who use the instruments on their jobs. This accounts for the fact that the N's range from 112 to 162. The figure indicates that the percentage of inspectors meeting the various standards ranged from a high of 66 per cent on the inside micrometer to a low of 9 per cent on the inside caliper in combination with the six-inch micrometer. The pattern of performance on the various vernier micrometers also seems significant. It will be noted that 43 per cent of those tested met the standard with the one-inch micrometer, 17 per cent with the two-inch, and only 11 per cent with the six-inch. The varying tolerances established for the instruments are the same as the tolerances that job analyses indicated had been established by the engineering department and are identical with those encountered in the shop.

After the completion of the study summarized, a similar investigation was conducted of forty-five toolmakers from another plant. Each toolmaker measured each of nineteen parts with a standard vernier micrometer. Each man measured each part independently five times

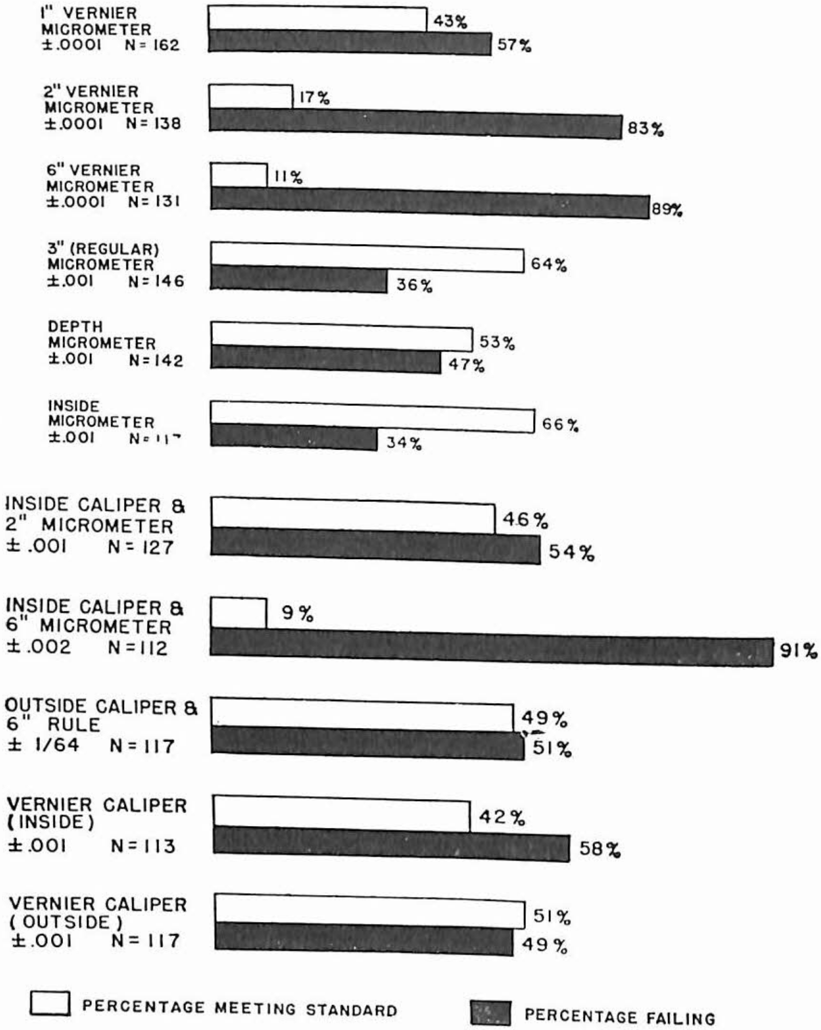


FIG. 13.3. The percentage of inspectors passing and failing various precision measuring instrument performance tests in an aircraft propeller plant. The open bars indicate the percentage meeting the standard and the solid bars indicate the percentage failing.

and the measurement recorded was the best judgment he could then make of the "true" measurement on the basis of these trials. After all of the readings by all men had been completed, the parts were measured with ultra-precision instruments and Johanssen blocks in order to obtain the closest possible approximations to the "true" dimensions against which to compare the measurements made by the men. A frequency distribution of the per cent of toolmakers who measured different numbers of the nineteen parts to within ±.0001 inch of the "true" dimension is shown in Fig. 13.4. The percentages of toolmakers who were able to measure various parts within specified tolerances were similar to those cited above for the inspectors. The results further showed that the larger the part, the less accurate the measurement.

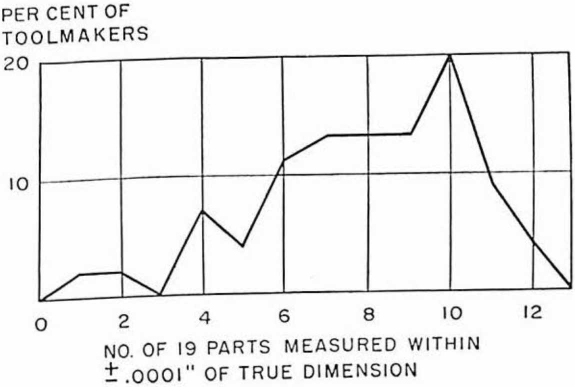


FIG. 13.4. Distribution of accuracy of toolmakers in measuring within ±.0001 of an inch with a standard vernier micrometer.

Another study of accuracy in using a micrometer was made by Evans.[3] Evans' results substantiate the findings of Lawshe and Tiffin[4] in that errors in micrometer measurements made by skilled workers are considerably larger than is assumed by most authorities in the field. He also found that the trade apprentices covered in his study were as accurate as journeymen, and that there was no significant relationship in the case of either group between accuracy and age, amount of experience with the company, or length of time on the present job.

Although studies such as those reported above suggest that people sometimes make poor inspectors, managements need not throw up their hands in despair. As indicated earlier, much can be done to improve the quality of inspection operations through proper placement, through the development of proper methods of inspection, and through training.

---

[3] R. N. Evans, "Training Improves Micrometer Accuracy," *Personnel Psychology*, 1951, 4, 231–242.

[4] Lawshe and Tiffin, *op. cit.*

The systematic investigation of an inspection problem relating to human beings usually requires the measurement of the inspection abilities of the inspectors on the job. The job sample tests discussed above are examples of this approach.

There are, however, other possible types of criteria of performance in addition to job samples. These include errors identified through reinspection, customer complaints (where they can be traced to the failure of individual inspectors to identify defective products), and length of job tenure. Although ratings can be used, they are likely to be related more with quantity of work inspected than with quality of inspection. The criterion of job performance that is used should be one that is appropriate, of course.

Having considered measures of job performance, let us now turn our attention to the various possible methods of reducing human errors in inspection processes.

PLACEMENT OF INSPECTORS

The development of standards for the selection and placement of inspectors involves essentially the same procedures as for any other job. We need to have, for individual inspectors in a group, a criterion of job performance (such as discussed above), and information about their personal characteristics (such as scores on various tests, biographical data, and so forth). It is then possible to analyze these two types of data to determine whether there are certain characteristics that, in general, differentiate between above-average and below-average inspectors. A few examples will be discussed.

Placement of Tin-Plate Inspectors. One of the purposes of the study of tin-plate inspectors was that of validating tests and biographical data for future use in the placement of individuals on this job. Toward this end, the 150 inspectors were given a battery of tests, these including a battery of visual skill tests,[5] the Minnesota Rate of

---

[5] The Keystone Telebinocular, Keystone View Co., Meadville, Pa.

Manipulations test,[6] the Purdue Hand Precision Test,[7] and measurements of reaction time and strength of grip. Information on height, weight, and age was also obtained. Table 13.2 summarizes the interrelationships of these tests and biographical-data items in terms of correlations between them.[8] This table also shows the correlations of the tests and biographical data with performance on the job sample test previously mentioned.

The first correlation shown in the table, it will be noted, is —.49, which indicates the relation between speed and accuracy on mixed plate. This figure means that there was a tendency for those who are above average in speed to be below average in accuracy, and vice versa. It will be noted further that the correlations between speed and accuracy for the various specific defects are all negative and that the amount of negative relation between speed and accuracy on off-weight plate is the smallest found in the group. The intercorrelations between accuracy for the various specific defects are all small, the largest being only +.35. This finding indicates that there is no decided tendency for girls who are above average in detecting one type of defect to be above average in detecting other types of defects.

The general conclusion that may be drawn from the correlations between the vision tests and the specific ways of scoring the coded stacks is that the visual acuity tests, distance and near, tend to pick the girls who are the most accurate on the job. The distance visual acuity test gave the highest positive correlations with accuracy in detecting all three appearance defects.

Table 13.2 also shows that the taller girls tended to be more accurate on the coded stack test, as indicated by the fact that all of the correlation between height and accuracy are positive. Likewise, in considering body weight we note that the heavier girls were somewhat slower but more accurate on everything but the off-weight sheets. The older girls, however, were not only slower but less accurate on everything but appearance defect 2. The more experienced girls were slightly faster but less accurate on nearly every defect.

Correlations between the remaining tests and the various ways of scoring the coded stack are given in the last four rows of Table

---

[6] Distributed by The Psychological Corporation, New York.

[7] Distributed by the Lafayette Instrument Co.

[8] The correlations in Table 13.2 are product moment correlations, computed as described in Appendix A, pp. 545–547.

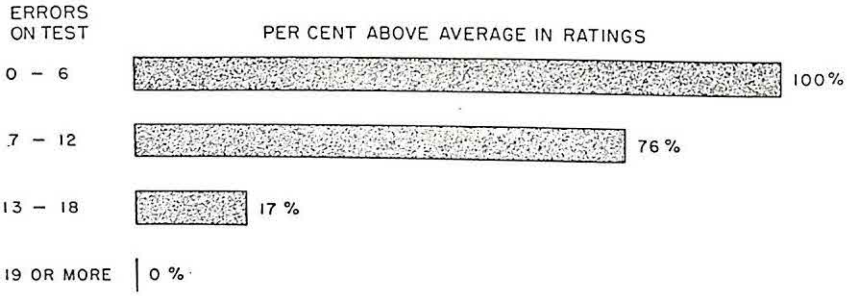


FIG. 13.5. Percentages of electrical inspectors with various error scores on perception test who were above and below average in job efficiency. (Adapted from Cook, op. cit.)

Placement of Bottle Selectors. In still another example, Coe[10] analyzed the relationship between biographical data and performance of a group of 65 bottle selectors. A bottle selector in a glass bottle plant is a person who inspects bottles for defects before shipment. Of 11 items of biographical information, it was found that 5 differentiated significantly between superior and other selectors. Performance was determined by a job sample test.

The results, given in Fig. 13.6, indicate that, generally speaking, superior selectors had had 10 or more years of education, were between 16 and 28 years of age, had no dependents, had two or more relatives who were also employed by the company, and weighed 128 pounds or less.

Although no selection and placement procedure is perfect, it can be seen by the above examples that systematic analysis can aid in identifying the characteristics of persons who will most probably be satisfactory on any given inspection job.

METHODS OF INSPECTION

Over a period of years there have been many improvements in instruments and other devices for use in inspection operations. Our concern here, however, is more that of improving methods of inspection from the point of view of the inspector, rather than that of improving instruments and devices. Here, again, various examples will be discussed.

---

[10] H. G. Coe, "Studies in Glass Bottle Selection" (Ph.D thesis, Purdue University, 1956).

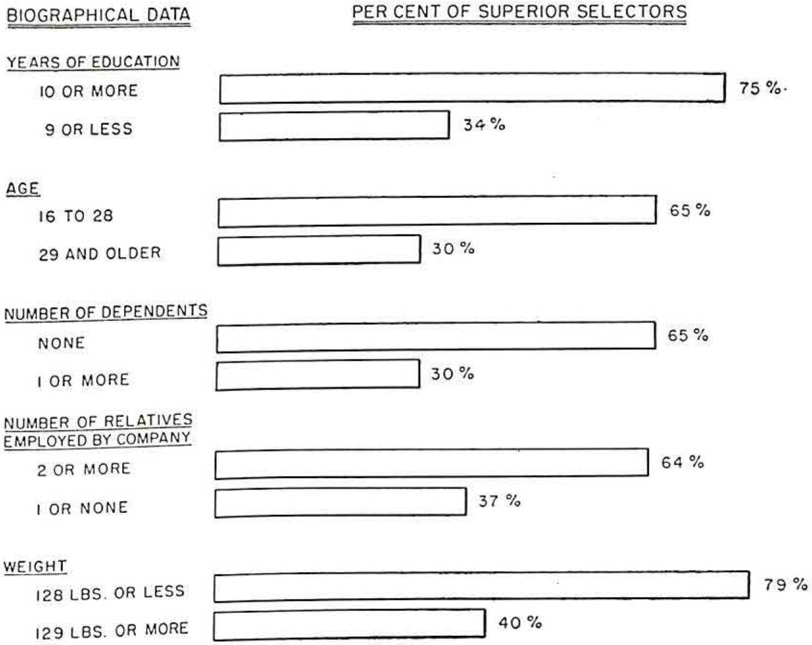


FIG. 13.6. Percentage of "superior" bottle selectors for various biographical data categories. (Adapted from Coe, op. cit.)

Methods of Inspecting Tin-Plate. The tin-plate inspection study, previously mentioned, had various objectives. One of these was that of developing the best method of inspection. For this purpose motion pictures were taken, at 1,000 frames per minute, of twelve inspectors selected from those who had taken the coded stack test. In this group were two inspectors who had been rated fast and accurate, two rated slow and accurate, two rated fast but inaccurate, two rated slow and inaccurate, and four rated average in both speed and accuracy according to the test-stack data. The pictures were taken in the shop under normal working conditions, and because of the general noise and shop activity, the operators were not aware of the exact moment at which the pictures were taken.

Obviously the pictures did not reveal the accuracy with which the inspector detected the defects. They did show, however, the normal speed of the operator, the relative body activity or exertion, the various methods used in grasping and moving the sheets, and the disturbance in the rhythm whenever a defective sheet was found. In general the speeds indicated by the analysis of these films correlated fairly well with the speeds indicated by the test stack.

The pictures were analyzed in an attempt to detect individual motion characteristics that might indicate causes for differences in accuracy. Since the detection of all but one class of defects depends upon vision, the eye fixations were located approximately in relation to the positions of the hands and the positions of the sheet.

In the normal handling of the sheets, the several stacks of sheets that have not been inspected, the sheets of first quality, those of second quality, and those containing various types of defects are aligned along a bench. The operator picks up a new sheet from the pile at her left and turns it over toward the right onto the adjacent pile of first-quality sheets where it normally remains. If a defect is discovered, the sheet is moved to the proper pile farther along the bench.

For most of the operators it was found that the eyes tend to follow the movements of the hands. Thus as the hands move toward the left to grasp a new sheet, the eyes attempt to scan the full top surface of the sheet within the very short period of time required for the grasp. As the sheet is picked up and turned over onto the adjacent stack, the eyes tend to follow the movement of the sheet that, in the upright position, presents only one edge toward the eyes. As the sheet approaches the right-hand stack, the eyes attempt to scan completely the second surface of the sheet; but as the release period is of very short duration, the scanning period is likewise very brief.

In this sequence, the eyes are attempting to see an object that is almost continually in motion, and it was found that the operators who most nearly followed this pattern were the lowest in accuracy on the test-stack data. This result is logical since it is more difficult to see a moving object and the continual attempt to do so reduces the accuracy of the operator.

On the other hand, the accuracy on the test-stack data was better for those inspectors whose eye movements were more nearly opposite the sequence described above. For example, the inspectors mentioned above, who were both fast and accurate according to the films and actual production records, used the following cycle of eye movements: as the hands moved to the left to pick up a new sheet, the eyes were scanning the second side of the previous sheet. The fact that this previous sheet was not in motion contributed to the ease of seeing possible defects, and the scanning period was increased to include nearly the full handling time of the next sheet. When the next sheet approached the second pile to the right of the operator, the eyes quickly shifted toward the left, to the first side of the top sheet on the new pile where the scanning period included the time for disposal of the previous sheet (see Fig. 13.7).

In this second cycle of eye movements, the eyes move quickly from one motionless object to the other and thus gain greatly increased scanning periods for detecting the defects. Accuracy data both from the test stacks and from production records correlated very closely with the eye-movement patterns of the twelve inspectors thus analyzed.

It has been demonstrated experimentally that the eyes can see a motionless object more accurately than one that is in motion, and the longer the scanning period, the greater the opportunity to observe defects. If the increased scanning periods are longer than necessary for effective inspection, the operator tends to handle the sheets faster. It is logical, therefore, that the girl who, accidentally or otherwise, has acquired this improved routine shall be both fast and accurate.

This cycle is opposite the more natural cycle of following the hands with the eyes; but as in many other operations, the natural way is not necessarily the best way. This less natural, but more effective, co-ordination of the eyes and hands is simple to learn, but because it is less natural, its technique must be explained and taught to the operator along with the other requirements of the job. This technique might be classed as one of the "tricks of the trade" not readily discovered without adequate instruction.

The Use of Examples as Standards. It has been rather clearly demonstrated in psychological investigations that people generally can make more accurate judgments when these judgments are *relative* rather than *absolute*. In other words, it is usually easier for people to judge amounts or degrees of qualities, traits, or characteristics of objects when *comparing* such objects with each other, than when attempting to judge such objects individually. This principle has applicability to inspection operations. The use of so-called "limit samples" illustrates the recognition of this principle. A limit sample is a sample of

Layout of Work Area. In some inspection processes the layout of the work area, or the manner in which the products are displayed to the inspectors, can have an important influence on the quality of inspection. A few examples will illustrate this.

Where products to be inspected are moving, as on a conveyor belt, there seems to be some evidence to support the contention that they should move laterally *past* the inspector rather than move *toward* the inspector. The desirability of this practice was indicated by the results of a study by Kephart and Besnard.[12] In their experiment, clear glass spheres were inspected for appearance under two test conditions—with the spheres rolling toward the inspector and with the spheres rolling laterally past the inspector. Half of the spheres used had been scratched. It was the job of the inspector to identify these "bad" spheres. The results showed significantly greater accuracy with the lateral view than with the end view. Although this particular inspection job would probably not often be encountered in exactly the same form, the basic point emphasized by the authors is that the *method* of investigating accuracy that may be expected with different work layouts while equipment is still in the design stage will clearly contribute to the most effective final layout of work to be performed.

Method of Display. An interesting method of investigation was that used by one company in connection with the job of inspecting roller bearings.[13] An eye camera was used to photograph the eye movements of 20 inspectors. On the basis of these films it was possible to determine the extent to which the inspectors used *foveal* vision or *peripheral* vision. A person uses *foveal* vision when gazing straight ahead at an object; the object is then usually in clear focus. Objects outside this central area of visual attention are seen by *peripheral* vision.

The analysis showed that the portion of the surface of the roller bearings that was typically seen by foveal vision varied for inspectors of different ability levels, as follows: best inspectors, 88 per cent; average inspectors, 68 per cent; and poorest inspectors, 32 per cent. It was found, further, that visual inspection was optimum when approximately five-eighths of the total surface of the rollers could be seen by foveal vision, and the remainder by peripheral vision.

---

[12] N. C. Kephart and G. G. Besnard, "Visual Differentiation of Moving Objects," *Journal of Applied Psychology*, 1950. 34, 50–53.

[13] "Eye Camera Shows How to Improve Inspection Technique," *Factory Management and Maintenance*, December 1953, 111, No. 12, 88–89.

Information such as that obtained in this investigation was then used in designing the machines that rolled the rollers past the inspectors in order to provide for the use of foveal vision for about five-eighths of the rollers' surfaces.

In another inspection operation, that of inspecting clear glass bulbs, it was found that a striped background served as an aid to the inspection.[14] By viewing the glass bulbs against a striped background, with adequate illumination, it was found that defects in the glass caused the straight background stripes to appear wavy.

It should again be pointed out that examples such as these are given *not* so much for their *specific results, as to illustrate a method;* in particular they indicate how some systematic analysis of inspection operations (or, for that matter, of other jobs) can aid in determining what work method or work layout would be best for the job in question.

TRAINING OF INSPECTORS

In the study on micrometer measurement by Evans[15] it was found that experienced toolmakers were no better in the accuracy of measurements than inexperienced students who had been given seven hours of training in the use of micrometers. This type of information argues strongly for the training of inspectors, since experience in an inspection job without training may not bring about a high level of inspection skill. Needless to say, the training should be specific to the job, and should be based on sound psychological principles of learning.

Training of Tin-Plate Inspectors. Since one of the purposes of the tin-plate study was related to training, it may be useful to discuss this study further, this time in a training context.

The results of the analyses of the job sample test of tin-plate inspection revealed several specific needs for training. In the first place, the accuracy of inspectors on the job was found to be far from the level of satisfactory performance that management considered desirable. Reference to Fig. 13.1 shows that, in the case of some of the defects, an appreciable number of the present employees were unable to spot over 10 or 20 per cent of the defective material.

---

[14] "Lighting Up for the Inspector," *Modern Industry,* December 15, 1951, 22, No. 6, 93.

[15] Evans, *op. cit.*

The training program was initiated with a series of inspectors' conferences. The inspectors were brought into conference in groups of approximately twenty, on company time. The first thing done in each of these conferences was to pass out a slip to each girl on which was printed her accuracy on the coded stack as a whole and on each of the four specific defects studied. This information gave each girl a clear indication of her weak as well as her strong points and resulted, in most cases, in a definite desire on her part to improve her level of performance. The fact that little correlation was revealed between accuracy on the different kinds of defects resulted in the fortunate circumstance that a girl who was very low in accuracy on one defect, or possibly even two defects, was likely to be average, or even above average, in accuracy on the other defects. This fact, coupled with the further fact that most of the girls were low in accuracy on at least one defect, prevented any employee from feeling that she had been singled out because of her low-quality work. Each inspector attended two of these training conferences, which lasted an hour and a half each and were separated by a period demonstration of the various kinds of defects found in tin plate.[16] The demonstration also included a review of the micromotion films and an explanation of differences between the patterns of eye movements with respect to the moving sheet. The films illustrated these eye-movement patterns and provided a means of teaching the most effective sequence of visual fixations. General improvement in performance following these conferences was noted.

The second innovation in training procedures that directly resulted from this investigation dealt with the training of new employees. The procedure followed in the past consisted in placing a new girl between two experienced inspectors. The expectation was that new employees would be able to ask questions of the more experienced employees. The actual results indicated that this procedure had not resulted in adequate training of new inspectors. Since hindsight is often better than foresight, it was easy to conclude after the investigation that experienced operators do not wish to be bothered with stopping frequently to answer the questions of a new girl. This knowledge resulted in each new employee finally setting up her own methods and standards of performance. It is not surprising that under such circumstances the standards varied markedly among the employees on the job.

---

[16] This procedure has been suggested by J. H. Mitchell, "Subjective Standards in Inspection for Appearance," *The Human Factor*, 1935, 9, 235–239.

The new procedure consisted in assigning an experienced and capable inspector as an instructor to each group of four new employees. The old employee was taken off the job of inspection, and since she had no assigned inspection work to perform she was able to devote all of her time to instructing the new employees. This procedure resulted not only in much more uniform standards of inspection among the newer employees than had heretofore existed but also in a marked reduction of the time necessary for bringing a new girl up to a satisfactory level of performance. As the newer girl became more adjusted to the job, less and less of the instructor's time was called for; and after a certain level of efficiency was reached, the instructor resumed her former work. Under this system the older employees assigned as instructors took considerable pride in the fact that they had been selected for the job of breaking in the newer inspectors, and the method was favorably received both by the older employees and the newer employees, as well as by the management.

SUMMARY

The general conclusions to be drawn from various studies of inspection jobs are that if adequate inspectors are to be obtained they must be selected with scientific precision and trained with corresponding throughness. No amount of desire on the part of a new employee to be a good inspector will result in satisfactory job performance if the employee does not have the specific qualifications for the job. And in like manner, no amount of qualification for the job will result in satisfactory performance when a systematic training program covering standards and procedures as well as "tricks of the trade" is lacking. In turn, it is also necessary that the methods of inspection be appropriate if the quality of inspection is to be satisfactory. A satisfactory inspection program depends upon these three elements: proper selection, appropriate method, and training. No one of these elements can be left out if the program is to meet high standards of inspection.

14

ACCIDENTS AND SAFETY

Accidents do not just happen. In dealing with them nothing is gained, except a temporary dodging of responsibility, by ascribing them to bad luck or chance. The concept of luck is used in explaining human behavior only when the really causative factors are unknown or when, for some reason, we have been unable to exercise adequate control over known factors. As those factors that are related to industrial accidents become more and more clear in the light of statistical investigation, the use of the concept of luck or chance becomes less and less necessary. In the light of such causative factors as are now known definitely to exist, Heinrich[1] has estimated that around 98 per cent of industrial accidents are preventable. Of these, around 90 per cent involve such things as faulty inspection, inability of the employee, poor discipline, lack of concentration, unsafe practice, and mental or physical unfitness for the job. These factors may be eliminated or greatly reduced by adequate supervision and proper job placement by employers. The importance of what might be called the psychological or personal factors in causing accidents is firmly attested by Heinrich's conclusion that only around 10 per cent of industrial accidents are due to distinctly physical causes such as faulty equipment or bad building conditions.

GENERAL CONSIDERATIONS

Psychological vs. Physical Conditions as Causes of Accidents. It is the purpose of this chapter to point out a number of essentially psychological factors that have been found by experimental and statistical investigation to be definitely related to industrial accidents. It is not meant, however, to imply that all accidents are related to, or can be explained by, the psychological factors involved. The importance of such external conditions as machinery safeguards and working conditions is fully attested by the favorable safety records that have so often followed the installation of a program that deals primarily with these factors. The present discussion is limited to the psychological aspects of safety, not because we fail to recognize the importance of the other aspects, but because this discussion is primarily intended to bring out the mental and personal features of a safety program.

---

[1] H. W. Heinrich, *Industrial Accident Prevention* (New York, McGraw-Hill Book Company, Inc., 1931).

Lost-Time Accidents, Home Cases, and First-Aid Cases. Any investigation intended to discover factors related to accidents must consider the severity of the accidents. One classification of accident experience makes use of three categories, namely:

1. Lost-time accidents—those in which the employee loses time from the plant in addition to the day, shift, or turn in which the accident occurred. Lost-time accidents involve the payment of industrial compensation.
2. Home cases—those in which the employee loses the remainder of the shift or turn on which the accident occurred. Accidents in this category do not involve the payment of industrial compensation.
3. First-aid cases—those in which the employee receives first-aid attention at the plant hospital or a first-aid station and then returns to his job.

One reason for classifying accidents according to a breakdown of this type is that the factors associated with a high frequency of minor accidents are not always associated with a high frequency of lost-time (or more serious) accidents. Table 14.1 clearly shows this situation for one year of accident records from a steel mill. For both types of accidents, the figures are given in terms of accidents per million man hours. A large difference in the relative frequency of accidents of the two types will be noted, the first-aid accidents far exceeding in number the lost-time accidents. A rather great difference in the order of importance or frequency of these accidents in the different departments will also be noted. For example, Department 7, which stands at the top in terms of first-aid accidents with 930 such accidents per

TABLE 14.1

Accident Data from a Steel Mill for One Fiscal Year

Department	*Lost-Time Accidents per Million Hours*	*First-Aid Accidents per Million Hours*
1	5.91	433
2	3.30	660
3	2.36	574
4	2.34	485
5	8.37	496
6	2.62	392
7	4.18	930
8	2.45	725
9	2.65	504
10	2.26	534
11	.50	242

million man hours for the year in question, is third from the top in lost-time accidents with 4.18 per million man hours. Department 5, which stands at the top in lost-time accidents with 8.37 per million man hours, is seventh from the top in first-aid accidents with 496 per million man hours. The rank order coefficient of correlation between frequency of these two types of accidents for the eleven departments studied was only .21. The conclusion indicated by this low correlation is that in this plant for the year studied only a very slight relationship exists between the hazard of a given department with respect to the possibility of lost-time accidents and the hazard of that same department with respect to more minor accidents. Thus, the factors that result in a large number of first-aid or minor accidents do not necessarily result in a correspondingly large number of lost-time accidents, and vice versa.

One should be cautious, however, in generalizing from a finding of this sort. In other plants, particularly those in different industries, a more marked relationship between frequency of minor and severe accidents may be found. In fact, this was reported by Wong and Hobbs[2] who concluded from their investigation that employees having the most frequent minor accidents also tend to have a disproportionate number of major accidents. The personnel research man working with a plant director of safety can readily determine the relation that exists between these two types of accidents for the plant under investigation. Since there was only a slight relationship between the two types of accidents in the steel mill where a considerable part of the data to be discussed in this chapter were obtained, care will be taken to identify in the case of each statistical summary whether the data refer to lost-time or minor accidents.

---

[2] W. A. Wong and G. E. Hobbs, "Personal Factors in Industrial Accidents—A Study of Accident Proneness in an Industrial Group," *Industrial Medicine*, 1949, **18**, 291–294.

The importance of reducing lost-time accidents is indicated by the fact that the plant from which a number of the investigations reported in this chapter were obtained has estimated that a lost-time accident involves a minimum cost of $2,000 to the company. First-aid injuries, on the other hand, are considerably less expensive. No actual figures on the exact cost of these latter injuries, however, are available.

Accident Proneness. When industrial safety first began to be given serious consideration by management, opinion differed considerably on the relative importance of *factors within the individual* and *working conditions* as causes of accidents. It is obvious that a safety program within a plant must depend upon specific knowledge of whether these factors are both operating, and, if they are, what their relative importance is for the plant in question. The concept of "accident proneness" is concerned with whether each employee tends to retain a given accident rate in comparison with other employees regardless of changes in the general physical condition of the machinery and working conditions within the plant. To the extent that each employee does tend to retain his same relative accident rate from one time to another, we may conclude that personal factors, peculiar to that individual, are affecting his accident record. This situation has been called *accident proneness.* Accident proneness is a descriptive rather than an explanatory term. When an individual employee is found consistently to experience more accidents than the average employee, he may properly be classified as an accident-prone employee. Such classification does not, however, explain *why* he is accident-prone; an explanation would require identification of the *causes* of his accident proneness.

It should be noticed in the above discussion that we have identified the accident-prone employee as one who *consistently* has more accidents than the average employee. An error has sometimes been made in accident studies when those employees who, during *one* period of time, have more accidents than the plant or department average are singled out as accident-prone employees. Mintz and Blum[3] have clearly shown that if only chance factors are causing the accidents, employees will not all be equally lucky or unlucky during a given period of work. If chance factors alone were operating, the situation would be quite analogous to tossing coins. If 1,000 men were each to toss ten coins, the average number of heads per man would be very close to five. But the individual men would vary in number of heads from very few (perhaps none) to many (perhaps ten). If we simply identify those men who tossed heads eight or more times, and call these men "heads-prone," we would be making the same error that is made when men who have more than the average number of accidents for one period of time are called accident-prone. But if such men are identified from accident records for a specified period of time and during a subsequent temporal period again experience more accidents than the average, we are reasonably safe in assuming that accident-proneness factors are operative. Maritz[4] has pointed out that the technique of correlating records for consecutive periods of time would seem to be indispensable in demonstrating the presence of accident-proneness factors.

The relative importance of working conditions and accident proneness as causative factors varies from one plant to another. In plants where the machinery used is unavoidably dangerous in itself, the former is of greater importance. In plants where the working conditions in and of themselves are not hazardous, the latter factor is of greater importance. When individual or personal factors are causing the trouble, it is clear that an adequate safety program must make provision for the individual handling of the high-accident-rate or accident-prone employees. It is, therefore, of vital importance for the safety director of a plant to know the extent to which his own employees are affected by accident-proneness factors.

In nearly every investigation of industrial accidents, accident proneness has been found to be a factor, and in some cases, a vitally important factor. The early work of Greenwood and Woods,[5] Marbe,[6] and others whose work has been summarized by Viteles[7] tends to support the statement that "accidents . . . happen frequently to some men and infrequently to others as a logical result of a combination of circumstances."[8] Several more recent investigations substantiate this general conclusion that, at least in many jobs, some employees tend to have consistently more accidents than can reasonably be accounted for by chance. The work of Slocombe,[9] for example, revealed a correlation of .51 between accidents for two successive years. Slocombe concluded, among other things, that the accident-prone men, as identified from their record for a given period of time, will subsequently experience more accidents than the average in 76 per cent of the cases. Mintz[10] has also reported that accident records of individuals have some validity as indications of future accident rates of at least some of those individuals, particularly those in which the persons have had a large number of accidents. It is not our intention here to resummarize the excellent digest of the literature on this subject already published by Viteles. Our purpose, rather, is to present certain statistics that have a direct bearing on this topic. These statistics both strengthen the case for the concept of accident proneness in industrial accidents and answer certain questions that previous investigations have not particularly considered.

---

[3] A. Mintz and M. L. Blum, "A Re-examination of the Accident Proneness Concept," *Journal of Applied Psychology*, 1949, 33, 195–211.

[4] J. S. Maritz, "On the Validity of Inference Drawn From the Fitting of Poisson and Negative Binomial Distributions to Observed Accident Data," *Psychological Bulletin*, 1950, 47, 434–443.

Table 14.2 summarizes certain facts pertaining to hospital visits from the employees of eleven departments of a sheet and tin mill employing around nine thousand employees for the years 1938 and 1939. The second, third, and fourth columns show, respectively, the average number of hospital visits per man for the year 1938, for the year 1939, and for the two years combined, for each of the eleven departments as well as for the mill as a whole (in the bottom row).

---

[5] M. Greenwood and H. M. Woods, "The Incidence of Industrial Accidents, with Special Reference to Multiple Accidents," *Industrial Fatigue Research Board, Report No. 4* (London, 1919).

[6] K. Marbe, *Praktische Psychologie der Unfälle und Betriebschäden* (München, 1926).

[7] M. S. Viteles, *Industrial Psychology* (New York, W. W. Norton & Company, Inc., 1932)

[8] *Ibid.*, 334.

[9] C. S. Slocombe, "Consistency of Operating Efficiency," *Personnal Journal*, 1930, **8**, 413–414.

[10] A. Mintz, "The Inference of Accident Liability from the Accident Record," *Journal of Applied Psychology*, 1954, **38**, 41–46.

TABLE 14.2

Frequency of Hospital Visits for Two Consecutive Years in Eleven Departments of a Steel Mill

Department	Freq. 1938	Freq. 1939	Freq. 38–39	Average Number of Hospital Visits in 1939 for Men With 0, 1, 2, etc., Visits in 1938									
				Number of Visits in 1938									
				0	1	2	3	4	5	6	7	8	9
1	.66	.85	.75	.61	1.05	1.54	1.14	1.64	2.20	5.00	2.50	....	....
2	1.13	1.31	1.22	.68	1.31	1.60	2.35	3.63	3.27	3.00	4.33	....	3.50
3	.81	1.03	.96	.63	1.14	1.56	2.16	2.56	4.50	3.00	1.00	....	7.00
4	.89	.81	.94	.55	.87	1.06	1.54	1.86	1.78	2.66	4.25	5.00	....
5	1.01	.99	1.00	.68	.95	1.00	1.94	2.23	1.60	3.33	5.00	....	....
6	1.20	1.27	1.23	.74	1.24	1.56	1.98	1.82	3.63	3.63	5.00	....	7.00
7	1.12	1.26	1.19	.81	1.15	1.45	2.00	2.29	3.44	3.21	11.00	4.00	5.00
8	1.32	1.21	1.26	.64	1.04	1.35	2.34	2.28	2.61	3.36	4.00	6.50	4.70
9	.71	1.05	.88	1.03	1.11	1.43	1.48	3.15	4.50	7.00	.....	....	....
10	.91	1.16	1.03	.53	1.26	2.05	2.41	3.18	3.50	4.25	4.50	7.00	8.00
11	.58	.88	.73	.71	1.22	1.47	1.14	3.00	5.50	....	.....	....	....
Average for All Depts.......	.94	1.07	1.01	.69	1.12	1.46	1.86	1.91	3.32	3.84	4.62	4.91	5.14

In the remaining columns of this table the average hospital visits per man for the year 1939 are given for the employees who had previously been classified according to the number of hospital visits experienced in 1938. Thus, for Department 1, the individuals with no hospital visits during 1938 averaged .61 hospital visits in 1939, whereas in this same department the employees with seven visits in 1938 averaged 2.50 visits in 1939. Although some inversions of the data from one category to the next occur, particularly in departments where the total number of men was not large, the general trend in all departments is toward an increase from left to right. When attention is turned to the data summarizing the mill as a whole, which are tabulated in the lower row, this trend is seen to be still more pronounced. For the mill as a whole, no inversions whatever are found in the tendency.

These data for the entire mill are shown graphically in Fig. 14.1 In this figure the number of hospital visits for 1938 is plotted along the base line or abscissa and the corresponding number of hospital visits for the following year along the ordinate. From these data we may confidently state that the average employee in this mill tends to retain approximately his same relative position with regard to hospital visits for any two successive years.

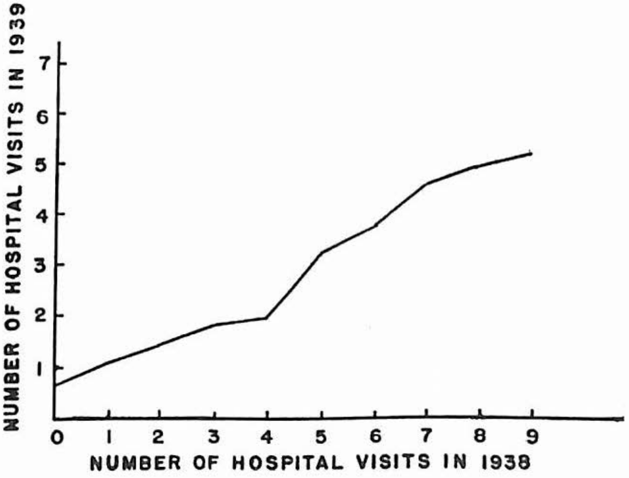


FIG. 14.1. Relation between number of hospital visits for two successive years among approximately 9,000 steel workers.

Accident Proneness vs. the Job in Relation to Accidents. The question may be raised why those employees who had nine hospital visits during the first year did not, on the average, have nine such visits during the second year instead of 5.14 visits as revealed by the data. If factors within the individual (accident proneness), were the sole cause of hospital visits, then the employees with nine hospital visits during 1938 should also have had on the average nine such visits during 1939. The discrepancy between the nine visits experienced by employees in this group during 1938 and the 5.14 average visits experienced by these same employees in 1939 may be taken as indicative of the relative relation of accident-proneness factors *vs.* external factors to hospital visits in the mill in question. In other words, we may say roughly that of every nine hospital visits made by the most accident-prone group, slightly more than five are essentially related to the accident proneness of these individuals and the remaining four cannot be related so clearly to individual factors within the employees. To prove the importance of the accident-proneness hypothesis, it is not necessary that the high-accident group for any year retain the same level of accidents for a subsequent year, but only that this group account during the subsequent year for a number of accidents significantly greater than the average of the plant. As revealed in Table 14.2, this situation is clearly the case for the mill in which these statistics were gathered.

An objection may be raised to these results as evidence of the fact than an individual or personal factor operates in the causation of

TABLE 14.3

Variation in Average Number of Hospital Visits per Year Among Employees on Eleven Different Jobs

Job	*Average number of hospital visits per year*
Craneman	3.55
Opener	3.54
Reckoner	2.96
Machinist helper	2.77
Leader	2.75
Sheet inspector	2.54
Shear helper	2.40
Assorter	2.36
Potman	2.10
Foreman	1.16
Roll turner	.47

accidents. This objection is that the job to which each employee is assigned may involve a specific accident hazard that, remaining the same or relatively the same from one year to another, results in a consistency of accidents for each employee from year to year, a consistency due more to the accident hazard of a given job than to any factors within the individual. It must indeed be admitted that each job has a relatively constant accident hazard and that this hazard clearly varies by rather marked amounts from one job to another. This fact is brought out in Table 14.3, which gives the average number of hospital visits per year for employees on eleven specific jobs in a steel mill. It will be noted that the average number of hospital visits varies from 3.55 in the case of cranemen down to .47 in the case of roll turners. It might reasonably be expected that this ratio of 7 to 1 found among this sampling of eleven jobs would be still larger if a larger sampling of jobs had been included in this survey. The table will suffice, however, to set forth clearly the fact that a specific hazard on different jobs does exist and that this hazard is likely to vary by a rather large amount from one job to another.

TABLE 14.4

Significant† Correlations between Personnel or Departmental Data and Accident Frequency and Severity

	Accident Frequency	*Accident Severity*
Per cent of employees who are male, production		.63
Per cent of employees who are male, salary		.50
Intra-company transfer mobility	−.44	
Sex ratio imbalance (men to women)		.51
Mean promotion possibility (mean ratings, twelve judges)	−.40	−.50
Per cent of suggestion quota met		−.54
Mean noise level	.42	
Youthfulness of employees (per cent under 26)		−.57
Tenure (per cent employed more than 12 months)		.55

Source: Kerr, *op. cit.*

† Less than one chance in twenty that they could have occurred by chance.

The fact that the accident hazard varies significantly among different departments of a plant is further revealed by a study made by Kerr.[11] His study dealt with fifty-three departments of the Camden Works of the Radio Corporation of America. Correlations were determined between both accident frequency and accident severity with forty types of personnel or departmental data. The correlations found to be significant are given in Table 14.4. Kerr suggests that a common explanatory factor of the relationships found might be *depressants to alertness*. But whether this is or is not a satisfactory explanation of the findings, the fact remains that significant relationships were found between the accident experience data and certain characteristics of the departments.

---

[11] W. A. Kerr, "Accident Proneness of Factory Departments," *Journal of Applied Psychology,* 1950, **34,** 167–170.

This situation makes it necessary to rule out the element of job differences in evaluating the concept of accident proneness. This was done in Table 14.5 by reassorting the data from the steel mill according to specific jobs. Table 14.5 is similar to Table 14.2 in general interpretation. In it, however, the employees are divided according to the particular job they are on rather than the department in which they are working, as was done in Table 14.2. Thus, all openers, regardless of the department in which they are working, are exposed to approximately the same accident hazards. It may be seen in Table 14.5 that for nearly every job the average number of hospital visits for 1939 shows a progressive increase. The combined data for the eleven jobs considered are summarized in the bottom row of Table 14.5 and are portrayed graphically in Fig. 14.2. We note nearly as

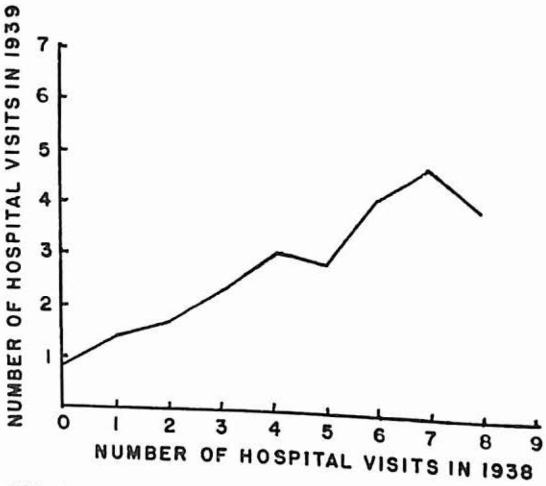


FIG. 14.2 Relation between number of hospital visits for two successive years among steel workers when hazard of the job is eliminated as a determining factor.

definite a relationship between accidents of two successive years in Table 14.2 and Fig. 14.2, when the element of job hazard has been essentially eliminated, as we previously noted in Table 14.2 and Fig. 14.1, when the job-hazard element might conceivably have had a significant effect upon the results. From the data, it seems safe to conclude that accident proneness, or individual accident susceptibility, is significantly related to the type of accident that results in hospital visits for the plant studied.

It is one thing to determine, as the data above demonstrate, that an appreciable number of hospital visits are due to accident proneness or factors within the individual. It is quite another thing, however, to identify the particular factors within the individual that determine accident proneness. In other words, once the concept of accident proneness has been established as an operating factor, the next question that arises is: For a particular job, what are the specific personal factors which differentiate between the consistent accident repeaters and employees with more favorable accident records? A discussion of this problem follows later in this chapter.

EXTERNAL FACTORS RELATED TO ACCIDENT PRONENESS

It is a recognized fact that many nonpsychological factors influence the accident rate and, to some extent, determine an employee's accident proneness. For example, it has been pointed out by Vernon[12] that the accident rate increases during the latter part of the working day. According to Vernon's results, this tendency is so marked that during a twelve-hour working day women experienced two and one-half times as many accidents as during a ten-hour day. Although fatigue has often been considered the cause of this increase, the fact that the time of greatest accident rate as compared with hours worked is reversed on the night shift indicates that psychological rather than physiological factors are operating. Although Vernon's results attach somewhat more importance to the length of the working day than do other investigations of this subject, it is quite commonly agreed that as the working day is lengthened, the accident rate increases in greater proportion than the increase in number of hours worked.

---

[12] H. M. Vernon, "An Experience of Munition Factories During the Great War," *Occupational Psychology*, 1940, **14**, 1–14.

TABLE 14.5

Accident Proneness for Men on Jobs with Similar Hazards

Job	Average Number of Hospital Visits in 1939 for Men with 0, 1, 2, etc. Visits in 1938									
	Number of Visits in 1938									
	0	1	2	3	4	5	6	7	8	9
Craneman	.46	.92	.78	2.20	1.00	1.00	3.00	....	....	....
Opener	1.00	1.54	1.50	1.83	2.33	2.00	6.00	7.00	2.00	5.00
Reckoner	.65	1.42	1.40	2.00	3.20	3.00	....	6.00	8.00	....
Machinist helper	1.19	1.36	3.67	3.75	3.00	3.66	....	....	....	....
Leader	1.38	1.38	2.00	3.00	6.00	....	....	....	....	....
Sheet inspector	.50	1.85	.78	....	....	....	....	....	....	....
Shear helper	1.63	1.40	1.72	2.25	....	5.00	....	....	....	....
Assorter	1.48	1.10	1.75	2.60	2.80	2.00	3.41	1.50	2.00	3.00
Potman	.40	3.00	1.30	....	....	....	....	....	....	....
Foreman	.48	1.05	1.42	1.08	2.75	....	....	....	....	....
Roll turner	.48	.45	....	....	....	....	....	....	....	....
Average for all	.85	1.40	1.63	2.33	3.01	2.78	4.14	4.83	4.00	4.00

TABLE 14.6

Departmental Differences in Hospital Visits Rate

Department	*Number of Men*	*Number of Accidents for 1939 & 1940*	*Accidents per Man per Year*
1	811	902	.55
2	573	1,144	1.26
3	480	723	.75
4	1,099	1,599	.73
5	336	555	.67
6	582	1,123	.96
7	624	1,238	.98
8	1,192	2,266	.94
9	373	529	.65
10	1,101	1,945	.88

The job itself, as we have already noted, may account for some of the variability in accident rates. Table 14.6 shows the hospital visits for ten departments of a sheet and tin mill. It will be noted in the last column of this table that the hospital visits per man per year vary from a minimum of .55 to a maximum of 1.26. This difference clearly reflects the hazard of working in the different departments. Several investigators have found a good many other nonpersonal factors, such as heat, ventilation, and humidity, to be related to industrial accidents. Osborne and Vernon[13] report, for example, that whenever the temperature increases or decreases above or below a normal and comfortable level of around 70 degrees, the accident rate becomes correspondingly higher.

It is also quite commonly accepted among safety men that an increase in the accident rate usually accompanies a step-up in production. Although this situation undoubtedly occurs in many individual plants when an attempt is made to step up production without adequate expansion of plant facilities, it is interesting to note that in a thorough study of the relationship between safety and production reported by the American Engineering Council[14] it was found, on a countrywide basis and over a period of years, that increased production in nearly every industry was accompanied by a decrease in the accident rate. This means that there is nothing inherently characteristic of a high production rate that tends to make automatically for a high accident rate. Rather it means that, as high production is achieved in the normal course of events by improved machinery and capacity for output, the accident rate will be likely to decrease along with the technological improvements and plant expansion. In any event, the effect upon safety of increased production, like the effect of temperature or job hazard, is essentially a physical rather than a psychological matter; and thus, important as this circumstance may be for the safety engineer who is interested in all angles of industrial safety, it is not primarily of psychological interest.

---

[13] E. E. Osborne and H. M. Vernon, "The Influence of Temperature and Other Conditions on the Frequency of Industrial Accidents," *Industrial Fatigue Research Board Report No. 19* (London, 1922).

[14] American Engineering Council, *Safety and Production: An Engineering and Statistical Study of the Relationship Between Industrial Safety and Production* (New York, Harper & Brothers, 1928).

PSYCHOLOGICAL FACTORS RELATED TO ACCIDENT PRONENESS

It has been implied in the above discussion that accidents are caused, and do not simply "happen." Basically, accidents probably can be attributed to either the *work situation* or to *personal variables* (accident proneness). The work situation generally determines to some degree the *liability* of accidents (some work situations are more hazardous than others). If, for people in the same work situation (and therefore subject to the same liability), there are consistent differences in accident frequency *among the individuals,* it can then be assumed that there are one or more *personal variables* that are contributing to the consistent individual differences. Knowing this, it is *then* in order to try to identify *what* personal variables are associated with accident frequency on the job in question. These will vary from job to job, and only through systematic investigation is it possible to identify *what* variable or variables are operating in the particular situation at hand. A few of these variables will be discussed.

Before discussing some of these variables, however, it should be added that an analysis of the *accident liability of the work situation,* and of any *personal variables* usually will not account for *all* of the variability in accident frequencies among individuals. That which we cannot explain, we attribute, then, to "chance." It should not be assumed, however, that if we cannot account for all accidents that the ones unaccounted for are really due to "chance"; rather, we should simply assume that we have *not yet accounted* for those factors that *are* causing them.

Vision. In spite of the fact that nearly every industry administers some form of vision test as a part of its employment procedure, vision surveys conducted among present employees in a number of industries reveal astounding differences in visual performance. These differences are due in part to the changes in vision that occur with increasing age and in part to the fact that an injury to, or deterioration of, an employee's vision after he has been employed often may go unnoticed unless an injury actually takes place on the company property during working hours. There are many reasons for conducting a periodic visual check-up of employees; one important reason is the relation that has been discovered between satisfactory vision and the frequency of accidents.

This problem has been investigated by comparing the accident experience of employees whose visual skills meet certain statistically determined standards with the corresponding accident experience of employees whose vision does not meet these standards. Table 14.7 summarizes data from several plants showing that employees whose

TABLE 14.7

Results of Visual Skill Tests in Relation to Accidents

N	*Job*	Percentage of employees who had not over two accidents during the period of the investigation among those who:	
		Pass Visual Requirements	*Fail Visual Requirements*
59	Passenger car drivers	71	42
116	Mobile equipment operators	74	71
65	Machine operators	67	50
15	Machine operators	67	44
29	Sheet metal workers	58	41
105	Maintenance men	81	75
63	City bus drivers	44	33
66	Intercity bus drivers	54	30
68	Supervisors	65	57
125	Machine operators	56	32
102	Laborers	82	89
15	Skilled tradesmen	45	33
828	All groups combined	65	57

vision is adequate for their jobs have fewer accidents than do employees with less adequate vision.[15]

For the fifty-nine passenger car drivers in the first job group, thirty-five had visual skills that met the recommended requirements. Of these thirty-five, twenty-five or 71 per cent showed a low frequency of accidents. On the other hand, among those passenger car drivers whose visual skills did not meet the standard requirements only 42 per cent showed a low frequency of accidents. A higher percentage of those workers whose vision was adequate for the job were in the low-accident-frequency group. Table 14.7 shows that for each of the twelve job groups similar results were obtained for all groups except laborers. In eleven of the twelve job groups, a higher percentage of workers whose visual skills meet the requirements for their jobs were in the low-frequency-of-accidents category than the corresponding percentage in this category among those whose visual skills did not satisfy the job requirements.

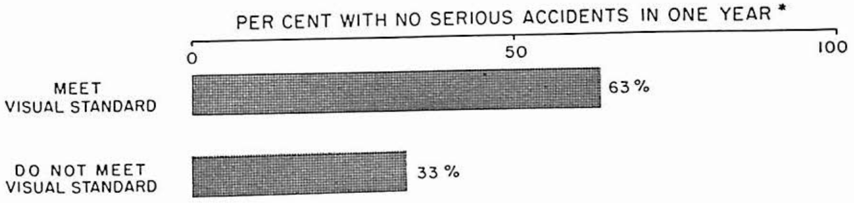
Of the total group of 828 workers, 463 or 56 per cent (not shown in Table 14.7) had visual skills that met the requirements of their job standard. Of these 463 workers possessing adequate visual skills, 65 per cent were in the low-accident-frequency group. Three hundred sixty-five or 44 per cent of the workers did not have the visual skills required by the job pattern. Of these 365 only 57 per cent were in the low-frequency-of-accident group. This is a difference of 8 per cent in favor of that group whose visual skills met the requirements of the standard. Statistical analysis (beyond the scope of this book) shows that the odds of these results occurring by chance are less than one in a hundred.

The results of one study of this sort are shown in Fig. 14.3. The visual standard for the job of paper machine operator was the one for machine operators as described on p. 153 and illustrated in Fig. 5.19 in Chapter 5.

The results of these several studies clearly reveal the relationship between vision and accident experience on some types of jobs. When these statistics were first shown to safety men from one of the plants in question, a thorough study was made of the individual case records of the various lost-time accidents. Nowhere in the description of these accidents, which had been written at the time of the accident

---

[15] N. C. Kephart and J. Tiffin, "Vision and Accident Experience," *National Safety News*, 1950, **62**, 90–91.



*A "serious" accident is one requiring the attention of the plant physician.

FIG. 14.3. Vision in relation to serious accidents among 104 paper machine operators. (Adapted from S. E. Wirt and H. H. Leedke, "Skillful Eyes Prevent Accidents," Annual News Letter, Industrial Nursing Section, National Safety Council, November 1945.)

by persons familiar with the background of each mishap, could there be found reference to the fact that faulty vision had played a part in causing the accident. These individual case studies did not reveal a deficiency of vision of the employee responsible for the accident. Yet the statistics show that the visual factor was operating. On an actuarial basis, it can be predicted that, if all plant employees had been required to pass the critical vision tests, the lost-time accident rate among the employees would have been markedly decreased. Uncovering by statistical investigation the importance of this factor, which had long escaped notice through ordinary case study of accidents, is a striking illustration of the value of the statistical approach in revealing hidden factors in any phase of human behavior.

No one should conclude from these statistics, however, that faulty vision is the *sole* cause of industrial accidents. Just as all kinds of factors may cause disease, so all kinds of factors may cause accidents. But the statistics do prove that faulty vision is *one* of the factors related to accidents in some circumstances. Certainly this factor is sufficiently important to deserve much more attention than it has received in many industrial plants in the past. Programs for such attention have been discussed in Chapter 5.

Age and Plant Service. A number of the investigations of the relationship between accident rates and plant experience or job experience of the employee have been reported. As might be expected, these investigations do not reveal entirely consistent results from one plant to another or from one industry to another. In most of these studies it has been found that the accident rate is higher among the younger, more inexperienced employees. This is the conclusion reached by Hewes,[16] Gates,[17] Schmitt,[18] Lipmann,[19] and Van Zelst.[20] Data obtained in a sheet and tin mill also show a drop-off in the accident rate with the age of the employees, their years of service in the plant, and their years of service on their present jobs. These relationships are shown graphically in Figs. 14.4, 14.5, and 14.6. Except for a low point of the curves for the very young or very inexperienced employees, there is a continuous decrease in hospital visits with increasing age or increasing plant or job experience.

Several possible explanations for the trends shown in Figs. 14.4, 14.5, and 14.6 have been offered. It has been suggested that the

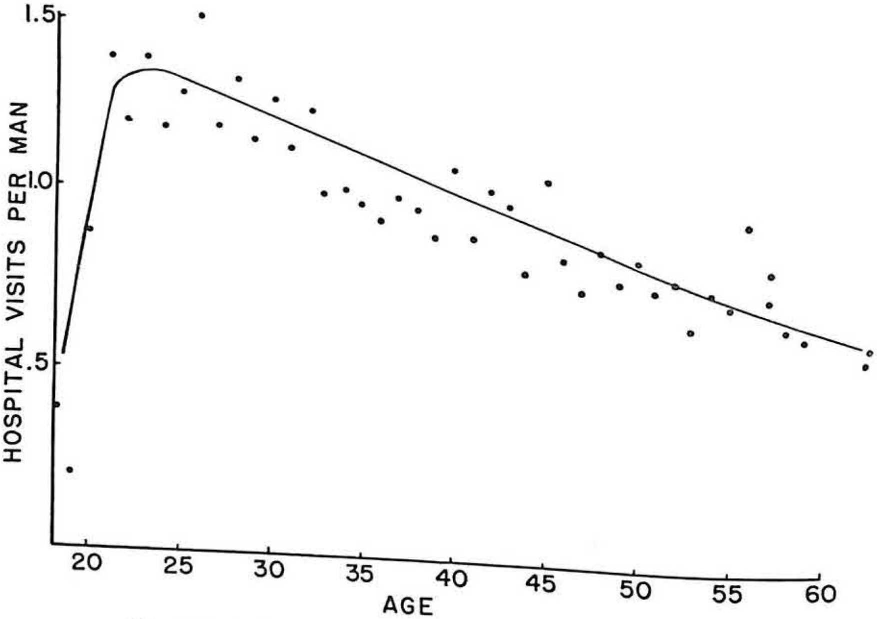


FIG. 14.4. Relation between age and hospital visits per man per year among 9,000 steel workers.

---

[16] A. Hewes, "Study of Accident Records in a Textile Mill," *Journal of Industrial Hygiene*, 1921, 3, 6.

[17] D. S. Gates, "A Statistical Study of Accidents in Cotton Mills, Print Works and Worsted Mills of a Textile Company," *Journal of Industrial Hygiene*, 1920, 2, 8.

[18] E. Schmitt, "Unfällaffinität und Psychotechnik im Eisenbahndienst," *Industrielle Psychotechnik*, 1926, 3, 144–153, 364–366.

[19] O. Lipmann, *Unfällursachen und Unfällbekampfung* (Berlin, 1925).

[20] R. H. Van Zelst, "The Effect of Age and Experience Upon Accident Rate," *Journal of Applied Psychology*, 1954, 38, 313–317.

younger employees are placed on the more hazardous jobs and that, as they become older and more "plant wise," they tend to be transferred to jobs of greater relative safety. It has also been suggested that the younger employee, who, in general, has relatively few family responsibilities, is less cautious and more likely to take chances than the older employee and therefore becomes involved in a greater proportion of industrial accidents. It has been further suggested that the younger employee is less acquainted with the possible dangers that exist in connection with the different machines and is thus more likely to subject himself to situations in which an accident is likely to occur. Any or all of these factors (and possibly other factors that have not been mentioned) may account for the results found. Whatever may be the cause or combination of causes, it is clear that in the plant studied the younger employee accounts for far more than his share of accidents. An adequate safety program must therefore take into account the age and service differential in accident susceptibility. This can be done by safeguarding the younger employee while he is becoming experienced on the job, by organizing safety conferences with particular reference to these employees, and by making other efforts to instill in this high-accident group those habits of work and methods of thought that tend to reduce the hazard.

It should be emphasized, however, that the relationships described above are not necessarily true of industry in general and that in certain other investigations a reversed type of relationship has been found. Stevens,[21] for example, found in the plants that he studied that frequency of accidents increased with age and that the mature workman is disabled for a longer period than is the younger man when an accident does occur. Shrosbree[22] also reports statistics from one industry to show that workers of longer service are a greater hazard. This is due, according to Shrosbree, to their becoming accustomed to danger and hence less careful than the less experienced employee. Although most studies show a decrease in accidents with age, the last two studies cited indicate that this is not invariably the case, and that, occasionally, even the reverse relationship may exist. The important thing to realize is not that a single rule necessarily

---

[21] A. F. Stevens, Jr., "Accidents of Older Workers: Relation of Age to Extent of Disability," *Personnel Journal,* 1929, **8,** 138–145.

[22] G. Shrosbree, "Relation of Accident Proneness to Length of Service," *Industrial Welfare,* 1933, **15,** 7–8.

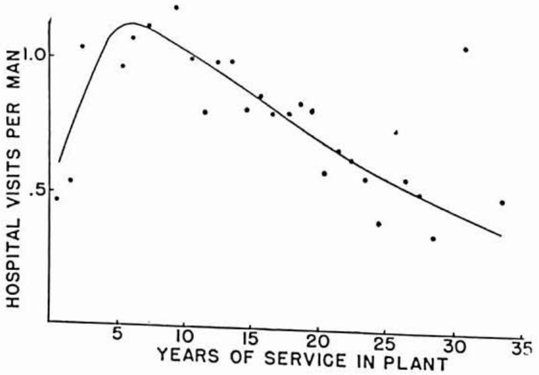


FIG. 14.5. Relation between years of service in the plant and hospital visits per man per year among 9,000 steel workers.

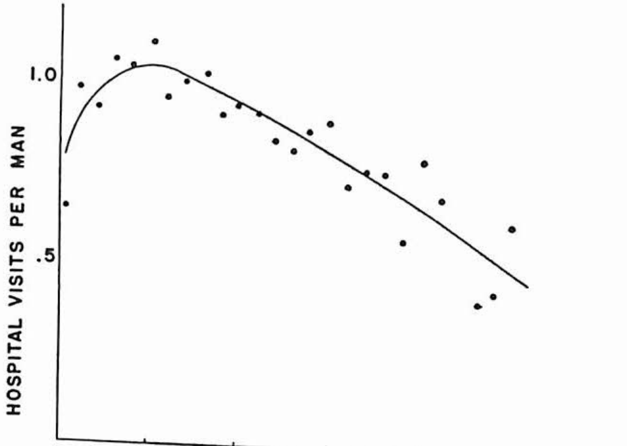


FIG. 14.6. Relation between years of service on present job and hospital visits per man per year among 9,000 steel workers.

applies to every industry, but that in any given industry the working conditions and other factors probably result in *some* sort of relationship. It is important for a safety director to know *what* relation exists in his plant and to direct his safety program specifically to that particular group of employees, young or old, who are working under the greatest accident hazard. Data of the type reported here are relatively easy to obtain in any industrial organization. As Howard[23] has emphasized, such data should definitely be in the hands of the safety director in formulating and administering his safety program.

Emotional Factors. Two emotional factors that have been found to be related to employee accidents in certain specific work situations are general emotional maturity and the emotional conditions at the time of the accident. An analysis of the causes of accident susceptibility among fifty motormen of the Cleveland Railway Company, published by the Metropolitan Life Insurance Company, is reproduced in Fig. 14.7. In this analysis four single items—faulty attitude, impulsiveness, nervousness and fear, and worry and depression—account together for 32 per cent of the accidents among the group studied. Since these four items are essentially emotional in nature, it may be inferred that emotional conditions account for a considerable proportion of accidents. It is well accepted psychologically that many individuals vary in their general emotional state between a "high" and a "low" condition. These conditions are often fairly regular in their occurrence and appear in a cycle that ranges from one extreme to another. This situation is readily recognized in its extreme form, as it results in the so-called manic-depressive psychosis. In a minor form, however, it characterizes many people who are in no sense of the word psychotic, but who do, nevertheless, experience definite "ups and downs" in their general emotional state. The relationship between these emotional cycles and the frequency of accidents has been shown in an investigation by Hersey,[24] who found that the average worker is emotionally low about 20 per cent of the time and that more than half of the four hundred minor accidents studied occurred during these low periods. According to chance alone only 20 per cent of the accidents would have occurred during such periods; the differential may therefore be taken as indicating the effect that an emotional depression has upon an employee's safety. Cardall also emphasizes the importance of emotional instability in causing accidents.[25]

---

[23] R. R. Howard, "Importance of the Human Factor in Industrial Accidents," *Journal of American Insurance*, 1942, 2, 11–14.

[24] R. B. Hersey, "Emotional Factors in Accidents," *Personnel Journal*, 1936, 15, 59–65.

[25] A. J. Cardall, "Psychological Factors in Accident Prevention," *Personnel Journal*, 1948, 26, 288–293.

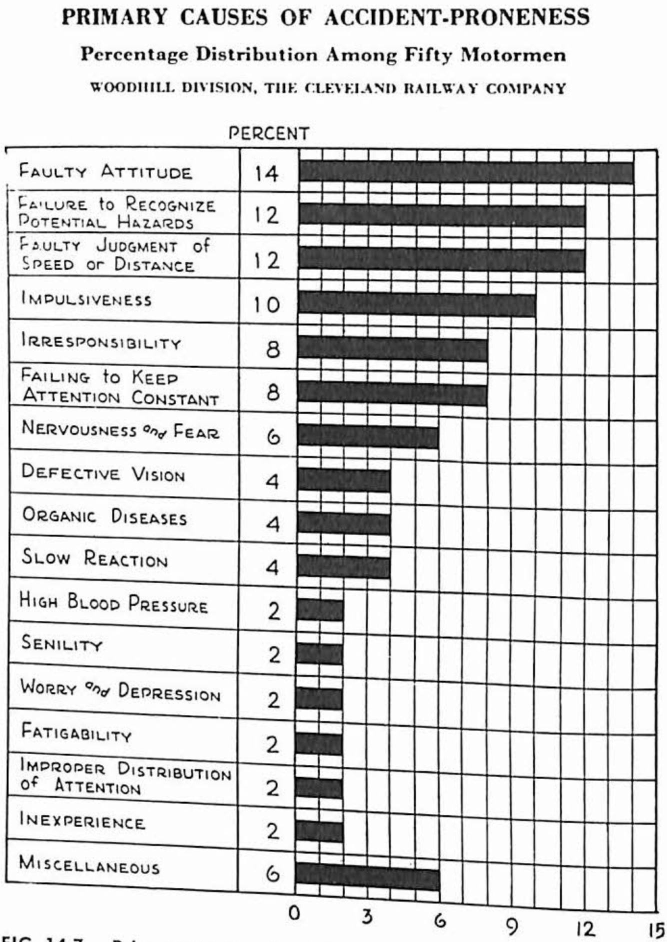


FIG. 14.7. Primary causes of accident proneness according to a survey among 50 motormen of the Cleveland Railway Company.

The general importance of the employee's emotional condition is further attested in a study by Hersey[26] in which it was reported that the production of industrial workers is around 8 per cent higher during periods when the men are elated, happy, hopeful, and cooperative than when they are suspicious, peevish, angry, disgusted, pessimistic, apprehensive, or worried. A favorable emotional condition thus not only is desirable from the standpoint of safety but also from the standpoint of plant productivity.

---

[26] R. B. Hersey, "Rates of Production and Emotional State," *Personnel Journal*, 1932, **10**, 355–364.

Supervisors and others dealing with human relations problems have long been aware that certain individuals vary from day to day in their agreeableness or ability to get along with others. Too often it has been the practice to handle such individuals either with a sharp reprimand or a general comment to the effect that so and so has a "grouch on" today. An employee who periodically gets a "grouch on" very likely is one who is afflicted with more than the average amount of cyclic fluctuation in his emotional state. To reprimand such an employee usually does about as much good as to reprimand a man with a broken arm for not doing his full share of work. Such an individual should be dealt with clinically to unearth the cause of his emotional condition. If this is impractical, he should be kept off a job that is in any way hazardous for the duration of his unsatisfactory emotional state. It should be emphasized that an employee who suffers from extreme variations in mood is not doing so for the fun of it. He cannot help himself and he, himself, is unhappier as a result of his condition than are any of the people who are in contact with him.

Aside from the fairly common fluctuations in temperament that are quite certain to be found in at least a few individuals in any large group of employees, there is the matter of general emotional maturity in responses that varies considerably from one person to another. It is entirely possible for an individual to be fifty years old by the calendar and yet be only five or six years old in terms of the level of his emotional behavior. A certain level of emotional development characterizes each state of the normal growth process. The baby cries and sleeps, the two-year-old kicks and screams, the eight-year-old fights, the adolescent sulks. The normal adult seldom indulges in any of these kinds of behavior; instead he has matured to a point where he recognizes and practices other, more social ways of gaining his end. Occasionally, however, we find an adult who has not grown up emotionally. He curses and fights like a child, sulks like an adolescent, and even occasionally cries like a baby. Such behavior obviously does not result in the degree of carefulness and responsibility that is necessary for safe practice in a modern industrial plant.

An indication of emotional characteristics of employees may often be obtained by means of personality tests such as those described in

Chapter 6. Such tests, however, are of little value in measuring personality or emotional characteristics of an individual who does not want to be measured. Unfortunately, an unwillingness to be measured is often found in the very employees who are most in need of emotional development. The unwillingness itself is often a symptom of emotional maladjustment. Such cases require definite therapeutic measures of the type described by Anderson[27] and Wright.[28] Anderson describes several points to be followed in making a psychiatric study of an employee who has had emotional difficulties. Wright describes in some detail the counseling plan that has been in operation for a period of years in several plants of the Western Electric Company. The essential feature of this plan is that provision is made for individual study and counseling of problem employees by professionally trained counselors. Sometimes a counselor will spend only a single period of a few minutes with an employee and will find that the problem of the employee has been satisfactorily solved. But often the counselor will find it necessary to devote a number of periods of one hour or even longer to a single employee before the trouble is really unearthed and the way made clear for a satisfactory solution.

Some type of individual study and treatment seems definitely to be necessary in the case of those employees who are handicapped by severe emotional problems. Such treatment, while clearly of personal value to the employee, need not be looked upon by management as a completely paternalistic venture; on the contrary, it should be viewed primarily as a financial and business venture. The solution of the emotional problems of employees is a paying proposition in terms of reduced accidents and increased production. However, employee counselling in industry is a specialized field and an adequate discussion of it is outside the scope of this book. The field is well covered by Cantor.[29]

Mental Ability. Numerous investigations have shown that for some jobs a certain minimum mental ability is required if an employee is to escape the hazards that occur on the job. In one situation, for example, Chambers[30] found that very few accident-prone individuals were above average in handwork, intelligence and learning ability, dependableness, and industry. The accidents for the most part happened to individuals who were relatively low in the traits associated with mental ability. In another investigation, Henig[31] found a definite relation between accidents in the Essex County Vocational School at West Orange, New Jersey, and scores on the Army Alpha Mental Test. The rapidly increasing use of some form of mental ability test as a standard part of employment procedure in many personnel offices is quite likely to be followed, as a secondary but very valuable result, by a reduction in accidents as well as an increase in general worker efficiency. The ever-increasing mechanization of modern industry is more and more increasing the importance of keeping away from the machine those individuals whose mental capacity renders them unable to recognize fully the dangers that are always present if unsafe practices are followed.

---

[27] V. V. Anderson, "The Problem Employee, His Study and Treatment," *Personnel Journal*, 1928, **7**, 203–225.

[28] H. A. Wright, "Personal Adjustment in Industry," *Occupations*, 1940, **18**, 501–505.

[29] N. Cantor, *Employee Counselling: A New Viewpoint in Industrial Psychology* (New York, McGraw-Hill Book Co., Inc., 1945).

In apparent contradiction to the results described above, a few investigators, for example Farmer and Chambers,[32] have found no correlation between degree of intelligence and accident frequency. The disagreement is more apparent than real, however, for investigations of this latter type have attempted to correlate the degree of intelligence with accidents among employees who are above a certain minimum intelligence level. In most work situations it seems clear that, above a minimum critical mental ability level, little, if any, relationship exists between further amounts of mental ability and freedom from accidents. But in many work situations it still remains quite necessary for an employee to possess this minimum amount of mental ability. Those who lack it are quite likely to be hazardous employees. It is primarily for the identification of this extremely low group that the use of mental ability tests is recommended as an important part of any accident-prevention program.

---

[30] E. G. Chambers, "A Preliminary Inquiry into the Part Played by Character and Temperament in Accident Causation," *Journal of Mental Science*, 1939, **85**, 115–118.

[31] M. S. Henig, "Intelligence and Safety," *Journal of Educational Research*, 1927, **16**, 81–87.

[32] E. Farmer and E. G. Chambers, "A Psychological Study of Individual Differences in Accident Rates," *Industrial Fatigue Research Board Report No. 38* (London, 1926).

stimulus of vision tests at the time of employment, many applicants would be quite willing to go to work with vision entirely inadequate for safe conduct within the plant.

Other tests dealing with such factors as mental ability and muscular speed may also be used to advantage, particularly in selecting employees for more hazardous jobs. The several investigations summarized on the relation between mental ability and accident experience all point to the general conclusion that a minimum score on an adequate mental ability test should be required of employees who are to be exposed to the normal hazards of a modern industrial plant. This does not mean that a particularly high critical score need, or should, be required for many jobs; but it does mean that a reasonable minimum score should be required and that an applicant who is so low in mental ability as to fall below that score should not be placed on hazardous jobs.

Safeguarding and Training Inexperienced Employees. The fact that young or inexperienced employees ordinarily experience more than their share of industrial accidents suggests the need for special care in building up a "safety consciousness" in the members of this group. This may be done in several ways. New employees often may be placed for a time on jobs that are not hazardous but that are close to those jobs where great care must be exercised. In this way the new employee will have an opportunity to learn by observation something about the machinery with which he later will be in contact. This procedure enabels the new employee to have a "breathing spell" during which he can adjust himself to the plant and overcome whatever feelings of uncertainty and anxiety he may have immediately after going to work in a strange (and, to him, often awesome) industrial plant.

Another method of helping the new employee is to assign him to periodic safety conferences. Conferences of this type are becoming more and more widely used as a tool of management, particularly on a supervisory level; more is said about the general nature and purposes of such employee conferences on p. 287. These conferences are of real value as one means of assisting the new employee in his task of becoming familiar with the hazards of his job. Every safety engineer has available a list of unsafe practices that apply to the several departments and operations of his plant. These unsafe practices are usually fairly well known to experienced employees who have seen men hurt when the safety rules were not followed, and who, therefore, are likely to avoid the unsafe practices with reasonable care. But the new employees, who often do not know what the unsafe practices are, are likely to endanger themselves not because they are essentially careless or foolhardy, but simply because they do not know *how* to be safe.

Walker and Potter[36] have emphasized the importance of worker participation in safety training. They point out that if employees are encouraged to participate in defining the safety requirements for jobs, the rules formulated will be accepted by them and provide a basis "for the constant retraining and reminding which the workers require" about safety.

Posters and placards may help the new employee to some extent, *if* the posters tell him specifically *what* to do and what *not* to do, and if he reads the posters and remembers what he reads. But too often posters simply show an accident and, in large type, say "Be Careful." Since only the maladjusted individual is likely to hurt himself voluntarily, the degree of carefulness of the ordinary employee is determined by (1) what he knows of safe and unsafe practices, and (2) how important he thinks it is to follow his knowledge on the subject. In other words, some employees follow unsafe practices because they actually do not know any better. The need in such cases is for specific information on what to do and what not to do. A systematic and comprehensive series of conferences covering this specific information is a much more effective means of informing the employees than is the chance and casual learning obtained from posters or from observations that employees may (or may not) make. Once the information has been given to the employees, still further education is necessary to make sure that they *believe* as well as *know* that certain practices are unsafe.

Personal Protection. The use of recommended safety measures, such as safety goggles and safety shoes, might be considered primarily as physical rather than psychological action if the use an employee makes of such safeguards were not so often determined by his own personal beliefs, attitudes, and, sometimes, prejudices. Many men believe that safety goggles are unnecessary even on jobs that are demonstrably dangerous because "they have never had an accident."

---

[36] W. S. Walker and C. J. Potter, "Worker Participation in Safety Through Job Analysis," *Personnel*, 1954, 31, 141–147.

This is one instance in which neither management nor employees can safely follow the adage that experience is the best teacher. Those who wait for experience to teach them that goggles should be worn to protect their eyes may have no eyes to protect after they have learned the lesson. The adequate utilization of safety measures requires not only that such measures be made available to the employees but also that the employees be educated in the use of these measures and *convinced* of their value. It is just as much a part of management's job to insure the success of the educational part of the program as to supply the safety measures.

Job analysis and accident records often reveal that employees on certain jobs are particularly liable to visual injury. Until such studies have been made, however, management is not always aware of the particular departments or jobs in which a visual hazard is most likely to exist. Table 14.8 shows a breakdown by departments of the relative frequency of seven different types of minor injuries experienced during two years by approximately 9,000 employees in a steel mill. It was noted that in one of the ten departments, Department 7, visual injuries were more frequently experienced than any other type of accident. In eight of the ten departments visual injuries were second in frequency, and in one department the visual injuries were third in order of frequency. For the plant as a whole, eye injuries were second in importance, being exceeded in frequency only by cuts and bruises. Many of the eye injuries reported were of course of a minor nature, perhaps involving little more than careful removal of a foreign body from the eye. Even injuries of this nature would be better prevented, however, for they involve not only loss of time for the employee and the company but also the possibility of serious complications requiring compensation and rehabilitation.

Influenced by such evidence as that presented in Table 14.8, many industries have adopted a mandatory safety-goggle program. Such a program requires the wearing of safety goggles by everyone in the plant at all times regardless of his job. The experience of organizations that have adopted such a program, such as the Pullman Company,[37] has been highly favorable. Indeed, these companies report a marked financial saving that is directly traceable to the reduction in eye accidents effected by the safety-goggle program.

---

[37] H. Guilbert, "No One Enters This Plant Without Goggles," *Factory Management and Maintenance,* 1938, **96,** 83–85.

TABLE 14.8

Departmental Differences in Type of Accidents
Data Based on Minor Accidents for One Year in a Steel Mill

Type of Injury	Department									
	1	2	3	4	5	6	7	8	9	10
Eye Injuries	14.4%	13.8%	14.1%	11.3%	25.5%	19.4%	38.8%	12.7%	20.4%	24.8%
Strains	3.5	3.3	2.9	2.6	3.4	2.6	3.7	3.2	3.5	2.7
Fractures	1.5	0.4	0.3	0.1	1.1	0.5	0.8	0.8	1.2	0.1
Heat	1.8	0.7	15.7	0.6	1.2	0.3	5.3	0.9	0.1	0.1
Burns	4.9	3.5	9.7	4.1	5.2	3.7	9.1	3.7	5.2	6.5
Cuts & Bruises	64.3	71.3	50.0	73.6	54.1	66.5	37.1	70.7	60.0	55.7
Miscellaneous	9.6	7.0	7.6	7.4	9.7	7.1	5.5	8.0	9.6	10.1
Total Number of Hospital Visits 1938–1939 (Minor Injuries)	1,217	1,390	861	1,852	652	1,416	1,433	3,000	657	2,269

Immediate Attention to Injury. Management finds it very much in order to be sure that every injury, no matter how slight, receives prompt and adequate attention. It is not enough to have hospital or first-aid service available for an employee to make use of if he wishes to do so. Means must be found to insure his making use of such facilities. Very often factors that have been totally overlooked by management effectively prevent many employees from taking advantage of available first-aid facilities. In one mill, for example, two machine shops contained identical machinery, employed approximately the same number of men, and turned out the same kind of work. An analysis of the accident records at the end of a year showed that the employees in one of these shops averaged 1.73 hospital visits per man for the year, whereas the employees in the other shop averaged only 1.10 hospital visits per man for the same period. So great a difference in the apparent hazards of the two shops was entirely unexplainable in terms of the men, the work, or the shops themselves. After thought was given to the problem, it was remembered that the hospital in which first aid and treatment was administered was next door to one of the machine shops but was about 500 yards from the other. The men in the shop close to the hospital took advantage of the service in nearly 60 per cent more cases of minor injury than did employees in the more distant shop. The solution to this problem is to educate the employees in the latter shop to walk 500 yards to receive treatment when treatment is needed, or to locate another emergency first-aid station closer to these men. Management should not and cannot take the point of view that the service is there and that employees who do not take advantage of it are on their own responsibility. Management must see that the service is utilized if a satisfactory safety record is to be obtained.

15

WORKING CONDITIONS

People's work performance depends upon several different *types* of variables. So far we have discussed *human characteristics* such as aptitudes, interests, and personality; *training; attitudes;* and the *motivation* of workers as resulting from different aspects of the work situation. Another type of variable that can influence employees' performance includes various aspects of the *working conditions.* We will consider here two primary types of working conditions—those relating to *physical environment* (illumination, noise, atmospheric conditions, and so on) and those relating to *time* (hours of work and rest pauses).

In order to assess the possible influence of such conditions on people in work situations we need to consider the *standards* by which variations in working conditions might be evaluated. This is the criterion problem presenting itself in another context. Taking, for example, variations in the work week, one can ask this question: "What criterion, or criteria, would be appropriate to use in determining an 'optimum' work week?" The same question can be posed for any other working condition.

CRITERIA IN EVALUATING WORKING CONDITIONS

Various types of criteria can be used as a basis for examining the influence of working conditions on people, the three primary ones being of a *physiological,* a *psychological,* and a *work-performance* nature. A few other criteria have also been used. The appropriateness of these will naturally depend upon the situation.

It is obvious that we are here dealing with some of the everyday experiences that people sometimes refer to as "fatigue" or "boredom."

Because of the rather ambiguous meanings of such terms, not only among laymen but also among research workers, in the following discussion of the various criteria we will endeavor to specify the meanings of terms as they will be used in this text.

Physiological Criteria. Human work is accomplished by certain physiological processes. As a person performs work, especially physical work, there typically are various changes in his physiological condition. If the work is severe enough or long enough an individual's physical ability to perform the task deteriorates.

The fact of reduction in physical ability to perform a task may be readily demonstrated by simple physiological experiments in which a muscle or muscle group is caused to undergo a simple rhythmic contraction. Thus, if one squeezes a coiled spring at intervals of two seconds he will find that his squeezes become less and less powerful until finally, if the task is maintained for sufficient time, only very small changes, or no changes at all, in the spring can be noticed. The energy used in the execution of a task of this sort, or in any muscular task, comes from potential energy that is stored in chemical form in the muscles. As this energy is expended, the muscles become less and less able to perform their task. This reduction in potential energy available in the muscles may be thought of as *physiological fatigue*. Starling[1] has stated that the phenomena of fatigue, from this point of view, probably depend upon two factors, namely, (1) the consumption of the contractile material or of the substances available for the supply of potential energy to this material, and (2) the accumulation of waste products of contraction. These waste products may be thought of as the chemical result of the muscular activity. Among them lactic acid is probably of great importance, because it is known that fatigue may be artificially induced in a muscle by feeding the muscle with a dilute solution of lactic acid.

Various physiological measures are used as indexes of physiological change in the organism. These include heart rate, blood pressure, oxygen consumption, breathing rate, blood composition, and electrical resistance in the skin. A measure that has been studied rather extensively by Brouha[2] is the heart rate recovery curve. This is obtained by counting the pulse rate at one-minute intervals during the first three minutes of the recovery period after the termination of work and while the subject is sitting quietly. The curve based on such readings indicates the actual value of the pulse and the rate of recovery toward resting level. Fig. 15.1 shows the heart rate recovery curves resulting from one, two, and three repetitions of a strenuous physical operation. The differences in the levels of these curves are indicative of differences in the physiological stress of performing the operation once, twice, and three times.

---

[1] E. H. Starling, *Principles of Human Physiology* (3rd ed; Philadelphia, Lea and Febiger, 1920), 209.

[2] L. Brouha, "Fatigue—Measuring and Reducing It," *Advanced Management*, January 1954, **19,** 9–19.

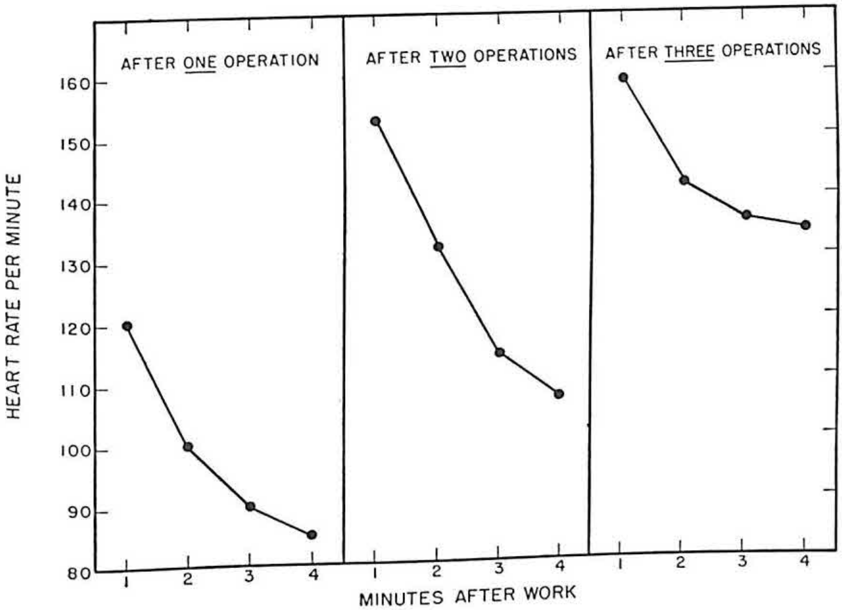


FIG. 15.1. Change in heart rate recovery curves after one, two, and three repetitions of a strenuous physical operation. (From Brouha, op. cit.)

Physiological changes are greatest, of course, with predominantly physical, as opposed to predominantly mental, work activities. On the basis of a review of various studies, however, Bills[3] points out that certain physiological measures are also practical in the study of mental work activities. In particular, he refers to the use of measures of muscular tension and electrical resistance in the skin. The physiological changes that occur in mental work are quite small, however, and to date have been of limited practical value as criteria of the effects of mental work.

---

[3] A. G. Bills, *The Psychology of Efficiency* (Harper & Brothers, 1943), p. 81.

While we use the term *physiological fatigue* to refer to physiological changes that reduce the ability of the organism (or a part of it) to perform, Bartley and Chute[4] use the term *impairment* in somewhat the same context.

Psychological Criteria. Entirely aside from the physiological changes that occur in connection with work (physiological fatigue), it is a matter of common experience that certain psychological changes sometimes occur during work, or when one is working under certain circumstances.

Although the dimensions and characteristics of such changes are not yet clearly delineated, there is reason to believe that there are at least two types of variables. One of these is the experience that we commonly call "boredom" and that usually is associated with work that is intrinsically uninteresting to the individual in question; most frequently boredom is associated with jobs that are repetitive or that are relatively simple. It should be pointed out, however, that it is the individual's *reaction* to the job that would cause him to describe a job as "boring"; no job is, or can be, intrinsically boring, since boredom is an attitudinal reaction of the individual.

The other type of psychological change that sometimes occurs in connection with work activities is what we commonly think of as "fatigue"—the feelings of tiredness, of weariness, of being "beat"—feelings which we all experience at one time or another. For purposes of convenience let us call this *subjective fatigue*. (It is also sometimes called *psychological fatigue*.)

Generally speaking, if we wish to find out how "bored" or "subjectively fatigued" a person is, we have to ask him. This, incidentally, raises some interesting problems in semantics, specifically in the meaning of words. Some words have different meanings to different people. Therefore, in expressing differing "degrees" of boredom or subjective fatigue one has to use words that have similar meanings to the person who is expressing his feelings and the person to whom he is expressing them.

---

[4] S. H. Bartley and E. Chute, *Fatigue and Impairment in Man* (McGraw-Hill Book Company, 1947), p. 48.

In this connection, McNelly[5] used the Thurstone scaling method (described previously in Chapter 10) in the scaling of words relating to subjective feelings of fatigue. A scale was constructed consisting of nine words which had been judged rather reliably in terms of subjective fatigue, and which represented varying "degrees." Actually, two forms were developed. One of the forms is given in Table 15.1. This shows each of the nine words or phrases, its scale value (actually the mean of the categories in which the several judges classified it), and its standard deviation (which shows the variability of those judgments). The results of an experimental laboratory study with this scale suggest that it may have some potential utility as a criterion of the feelings of fatigue that people experience.

Work Output. Still a third type of criterion that has been used in studying the influence of various working conditions on people is their *work output*, or some other appropriate indication of their job performance. In some studies, including those of Bills,[6] work decrement has been used, this being the reduction in work output from one time to another, or from one condition to another.

TABLE 15.1

Words and Phrases in a Scale of "Subjective Fatigue" showing Mean Scale Values and Standard Deviations of Judgments

Word or phrase	*Mean scale value*	*Standard deviation*	*Rounded scale value*
About to fall over	1.1	.30	1
Fagged	2.2	.36	2
Let down	3.2	.51	3
A little tired	3.8	.40	4
Average	5.0	.00	5
Fairly well	5.9	.57	6
In gear	7.0	.63	7
Very good	7.9	.54	8
Terrific	8.9	.30	9

Source: McNelly, *op. cit.*

Where work output tends to diminish, the term *industrial fatigue* is sometimes used. Vernon[7] has defined industrial fatigue as "the sum of the results of activities which show themselves in a diminished capacity for doing work." Since, however, in comparing the effects of various working conditions one may find that *one* condition results in *greater* output than *another*, the term *work output* is somewhat more flexible than *industrial fatigue*, which, by definition, implies a decrement only.

---

[5] G. W. McNelly, *The Development and Laboratory-Validation of a Subjective Fatigue Scale* (Ph.D. thesis, Purdue University, August 1954).

[6] A. G. Bills, *op. cit.*

[7] H. M. Vernon, "Industrial Fatigue in Relation to Atmospheric Conditions," *Physiological Review*, 1928, 8, 130–150.

Other Criteria. There are other criteria that have been used, or proposed for use, in studies of the effects of work.

Lauru platform. A very interesting and promising criterion is one recently developed by Lauru,[8] a French physiologist. He has developed a method for recording and measuring the physical forces exerted by the body. This is done by the use of an equilateral triangle platform (about 3 feet on each side) on which a subject stands while at work. Underneath the platform are five piezo-electric quartz crystals, connected to an electronic system. As the subject stands on the platform performing a task, his center of gravity shifts one way or another. The physical forces created by this shift in gravity cause the platform to tilt very slightly, thereby compressing certain of the quartz crystals and creating changes in electrical voltage which are picked up and amplified by a multichannel recording instrument. The system is responsive to changes in forces covering a wide range. The device records forces in three directions: frontal (from front to back), transverse (from side to side), and vertical (up and down).

Lauru claims that the equipment, by measuring physical forces of the body, actually provides an index of physiological stresses. In a sense, then, this device indirectly measures physiological changes. An application of this approach will be discussed in the next chapter.

Efficiency of work. Aside from the specific types of criteria discussed above, Ryan[9] proposes a concept of *efficiency* that incorporates various elements involved in work. Borrowing from physics and mathematics, he considers efficiency to be a ratio, as follows:

$$\text{Efficiency} = \frac{\text{Output}}{\text{Input}}$$

In this equation, *output* is considered to include not only any tangible products of work, but also any pleasure or satisfaction that the worker may derive from his accomplishment. *Input*, in turn, includes a complex set of factors, in particular any factors that may be thought of as *costs* of work to the worker; among these are physical energy required, the time of the individual, the effects of the work on health, on his personal adjustment in society, and on his ability to enjoy leisure time and recreation.

---

[8] L. Lauru, "Physiological Study of Motions," edited and translated by L. Brouha, *Advanced Management*, March 1957, 22, 17–24.

[9] T. A. Ryan, *Work and Effort* (Ronald Press, 1947), p. 20.

In proposing this concept, Ryan points out that work output (even when it is high), is not necessarily an indication of the efficiency of an individual, since the "input" (cost) of performance may be disproportionately high, thereby actually resulting in a low efficiency index.

Ryan recognizes the present problems of measuring some of the elements in this ratio, but suggests that research should be directed toward making this possible.

Relationships among Criteria. For one speculating upon the subject of "fatigue" and related concepts, it is a temptation to look upon two aspects of the phenomenon—the physiological and the psychological—as different aspects of the same fundamental change. According to this point of view, feelings of boredom and tiredness occur in proportion as physiological changes have taken place; and vice versa, actual physiological changes have taken place to the extent that feelings of tiredness or boredom are experienced by the subject.

If this simple relation were true, the problems of fatigue and its elimination would be greatly simplified. Unfortunately, numerous physiological and psychological experiments have shown that no such simple relationship exists. Indeed, it is necessary to look only at one's own experiences to understand that exceptions to such a relationship are by no means uncommon. Anyone who has driven a car a very long distance to a vacation resort knows that interest in reaching the resort is often maintained long after the body, from the physiological viewpoint, would rather sleep.

Anyone who has become "bored" by his job knows that such feelings of boredom often become prominent long before an excessive amount of work in the physiological sense has been performed. Further, a person may feel subjectively tired, but still maintain an intense interest in an activity—such as a game of golf. In still a different situation, a person may feel tired and bored with his work, and even be physiologically fatigued—at least to a moderate degree—and still maintain an appreciable level of work performance, presumably because of the strength of the incentives. In this connection Arps[10] has pointed out that a production curve can be maintained at a fairly high and constant level if incentives are sufficiently strong, even though physiological changes indicate a great or even unwise expenditure of energy.

Considering the three major types of criteria with which we are dealing (physiological changes, psychological changes, and work output), it can be seen that they can, and often do, occur independently. While the magnitude of the correlations among them undoubtedly depends much upon the type of work in question, the correlations generally will probably be found to be moderate or low. There is some reason to believe, however, that for many types of work the psychological variables (such as boredom) are more markedly related to employee production than are physical or organic conditions. In one series of studies it was found that "fatigue" usually implies dissatisfaction and that physical conditions varying within ordinary limits have an insignificant effect upon the efficiency of the worker.[11] In another investigation it was found that whatever is done by management to attract the workers' interest or indicate a concern for their welfare decreases "fatigue" and increases output.[12]

Physiological fatigue (reduced physiological ability to perform) is a much more important variable in heavy manual types of jobs, where the reduced ability may actually bring about a reduced output, than it is in light or sedentary work.

Discussion of Criteria. The type of criterion used in any study or experiment naturally depends in large part on the nature and purpose of the investigation. For example, in studying the effects of temperature on people, one might appropriately use some measure of physiological change. If, on the other hand, one were investigating rest pauses for office workers, it might be more in order to use some measure of work output.

There are, however, certain considerations that frequently make work-output measures the most feasible to use in many practical circumstances. In the first place, the measurements of physiological changes during work are relatively difficult to make and, where used, frequently are time-consuming and costly. Developments are taking place, however, that may make such criteria more feasible. The research of Brouha[13] and the work with the Lauru platform[14] (which is associated with physiological stress), are examples of such developments. Some studies have been carried out using scales of subjective fatigue and boredom, but these presently cannot be used on a broad scale.

---

[10] G. F. Arps, "A Preliminary Report of 'Work with Knowledge *vs.* Work without Knowledge of Results,'" *Psychological Review*, 1917, **24**, 449–455.

[11] Committee on Work in Industry of the National Research Council, *Fatigue of Workers and Its Relation to Industrial Production* (New York, Reinhold Publishing Corporation, 1941).

[12] R. H. Flinn, "Fatigue and War Production," *Medical Clinics of North America*, 1942, **26**, 1121–1143.

Further, in the final analysis, a business or industrial organization must operate in the black in order to continue to exist. One that operates in the red for very long will go out of business. This economic emphasis naturally supports the use of a work-output criterion in considering the effects of work environment upon people. Although this emphasis on an output type of criterion does not in any way preclude the concurrent use of physiological and psychological variables, it is obvious that in practical situations one simply cannot disregard the need for maintaining reasonably adequate output levels.

Moreover, granting that the correlations among physiological, psychological, and output variables probably are not very high, they probably suggest that the working conditions that contribute to the minimizing of physiological and psychological costs in performing work probably also contribute in the direction of optimizing work output. For example, the creation of atmospheric conditions that reduce physiological costs probably also would make possible the maintenance of a higher output level. Similarly, if occasional rest pauses tend to reduce the "boredom" of work, the result may be one of at least maintaining, and possibly one of increasing, total output.

ILLUMINATION

Illumination is one of the most important aspects of the work environment for people on many types of jobs.

Illumination and Visual Acuity. In a number of investigations such as those reported in Chapter 5 it has been found that an employee's job performance is related to his visual acuity at the work distance. Visual acuity, however, is dependent not only upon the visual mechanism of the employee but also upon the illumination under which the work is done. The relation between visual acuity and illumination has been plotted by Troland[15] according to data published by König.[16] The results plotted by Troland show that acuity increases over a considerable range of variation of intensity in almost direct proportion to the amount of illumination. These results suggest that employees on certain types of industrial work should be provided with sufficient illumination to insure a satisfactory level of visual acuity. Wetzel[17] has pointed out that gains in visual acuity are very perceptible up to about 40 foot candles.

---

[13] Brouha, *op. cit.*

[14] Lauru, *op. cit.*

A study by Kuntz and Sleight[18] indicates that persons with *subnormal* vision benefit *more* in ability to see by increasing illumination levels than do persons with *normal* vision. This probably suggests that where persons with subnormal vision are used on jobs that involve visual discriminations of some consequence, it is particularly important to have adequate illumination.

Adequate vision for a job can, in part, be achieved by proper placement (placing people with adequate vision), and by proper visual care (such as using glasses if a person's visual skills are inadequate without a correction).

Illumination and Brightness Contrast. The ability to make visual discriminations, however, depends in part on still other factors, among them *brightness contrast*. This is the relationship of the contrast between the detail of objects to be discriminated and their background, such as the contrast of the black letters on this white paper, or the defects in a piece of cloth that an inspector would have to identify.

When there is not much brightness contrast, the illumination needs to be greater, for equal visibility, than when contrast is great. This is shown by the results of an experiment by Luckiesh[19] in which three levels of illumination (10, 20, and 100 foot candles) were used with visual targets of two levels of brightness contrast (black on white, and black on gray). The time required for perfect identification was determined, the results being shown in Fig. 15.2. It will be noted that for each level of illumination the black object was seen more quickly on the white background; but further, it can be seen that by increasing the illumination, the black objects on gray could be identified in shorter and shorter times.

---

[15] L. T. Troland, *The Principles of Psychophysiology*, II, "*Sensation*" (New York, D. Van Nostrand Company, Inc., 1930), 86.

[16] A. König, "Abhangigkeit der Sehschärfe von der Beleuchtungsintensität." *Sitzungsber. der Akad. der Wiss., Berlin, Bd.* 1897, **13**, 559–575.

[17] M. Wetzel, "L'eclairage dans l'industrie," *Recherches et inventions*, 1927, **8**, 81–95.

[18] J. E. Kuntz and R. B. Sleight, "Effect of Target Brightness on 'Normal' and 'Subnormal' Visual Acuity," *Journal of Applied Psychology*, 1949, **33**, 83–91.

[19] M. Luckiesh, *Light, Vision, and Seeing* (New York, D. Van Nostrand Company, Inc., 1944).

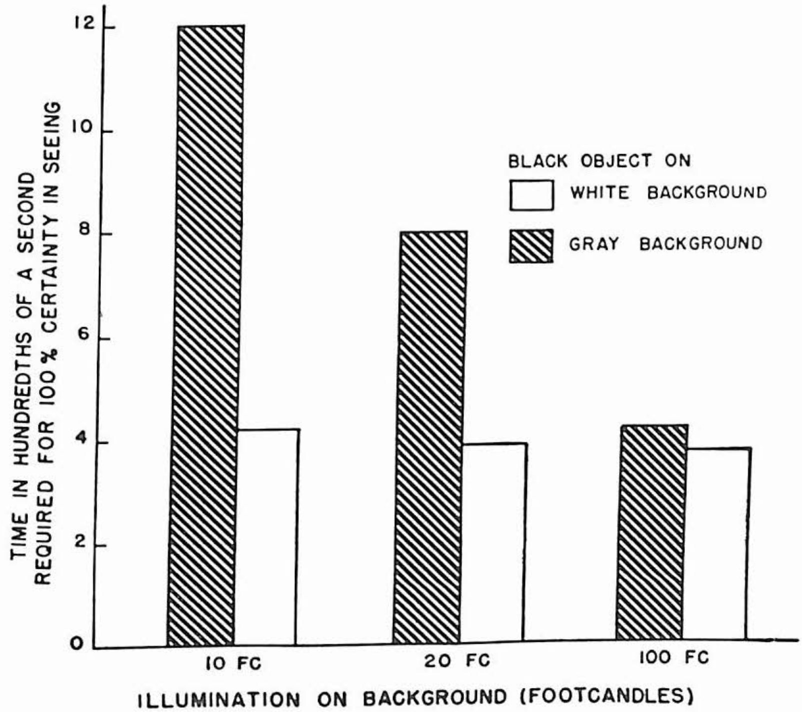


FIG. 15.2. Relation between brightness contrast and certainty of seeing. (Adapted from Mathew Luckiesh, Light, Vision, and Seeing, New York, D. Van Nostrand, 1944, p. 132.)

Optimum Illumination Levels. It can readily be understood that the optimum level of illumination varies for different types of work. In attempting to determine the "optimum" levels for various jobs, however, it should be pointed out that the experiments conducted in this field, and particularly the interpretations drawn from them, have resulted in rather marked disagreement.

It is not our purpose to review the details of this controversy, but it should be pointed out that the optimum levels of illumination for various tasks that have been recommended by Luckiesh and by Luckiesh and Moss[20] are consistently higher than those recommended by Tinker, Hoffman, and Birren.[21] The differences are illustrated by the fact that Luckiesh[22] recommends 100 foot candles or more for such tasks as fine needle and pen work, 50–100 foot candles for proof reading, 20–50 candles for clerical work, 10–20 foot candles for ordinary reading, 5–10 foot candles for visual work that is more or less interrupted or casual, and 1–5 foot candles only for the perception of large objects. Tinker,[23] on the other hand, states that 40–50 foot candles will be adequate for the most severe tasks found in work-day situations, 30–40 foot candles for tasks comparable to discrimination of 6-point type, 20–30 foot candles for situations that involve reading handwriting and other comparable tasks, 15–20 foot candles for situations such as reading newsprint, and 10–15 foot candles for reading good-sized print (10- to 11-point) on a good-quality paper.

The major reason for the disagreement in recommended levels of illumination may be found in differences in the criterion used in evaluating the "best" level. Luckiesh has used preferences of the worker (or reader), visual acuity, visibility measurements, nervous muscular tension, frequency of blinking, and change in heart rate. Every one of these criteria has been criticized by other experimenters as being invalid for this particular purpose. The literature of this controversy has been summarized by Tinker.[24] The majority of other persons who have worked on the problem or have attempted to interpret the work of others seem to favor the lower levels of illumination that Tinker recommends.

---

[20] M. Luckiesh, *Light, Vision, and Seeing* (New York, D. Van Nostrand Company, Inc., 1944). M. Luckiesh and F. K. Moss, *The Science of Seeing* (New York, D. Van Nostrand Company, Inc., 1937).

[21] M. A. Tinker, "Illumination Standards for Effective and Easy Seeing," *Psychological Bulletin*, 1947, **44**, 435–450. See also, A. C. Hoffman, "Luckiesh and Moss on Reading Illumination," *Journal of Applied Psychology*, 1947, **31**, 44–53, and F. Birren, "Light Control *vs.* Light Intensity," *Transactions of the American Academy of Opthalmology and Otolaryngology*, 1950, **54**, 590–595.

[22] Luckiesh, *op. cit.*

[23] Tinker, *op. cit.*

[24] *Ibid.*

ATMOSPHERIC CONDITIONS

Different kinds of variables can be thought of as being aspects of our atmosphere. Besides temperature and humidity there are also the factors of air flow, barometric pressure, composition of the atmosphere, and sometimes toxic conditions. There is also the factor of temperature of objects in the environment; this is not strictly an "atmospheric" condition but it certainly relates to this subject. Our discussion will deal largely with the more common aspects of the atmosphere, especially temperature and humidity.

The Heat-Exchange Processes. The metabolic processes of the body result in the generation of heat, some of which the body normally has to dispose of. In disposing of such heat, the body is continually attempting to maintain a thermal equilibrium with its environment. The heat-exchange process is the method by which the body attempts to achieve this balance.

There are four ways in which this heat exchange takes place. Convection is the transmission of heat by a fluid that occurs when there is a temperature difference between the object and the gas. In the case of people, the body typically transmits heat to the air (which is technically a fluid), although when the air temperature gets above body temperature the transmission is reversed. *Evaporation* is another method of heat exchange; this consists primarily of evaporation of perspiration, and to some degree of vapor that is exhaled from the lungs in breathing. *Radiation* is the process of transmission of thermal energy between objects such as between the sun and the earth. Such transmission occurs either with or without an atmosphere if there is a difference between the temperatures of two objects, the warmer object losing heat to the cooler. Usually people transmit heat to *other* objects by this method, but occasionally a person may be in a situation where the objects transmit heat to him, such as from boilers or heated metal. *Conduction* is the transmission of heat by direct contact, such as with chairs or the floor. Our clothing usually so insulates us that this is a very unimportant method of heat exchange.

Conditions That Affect Heat Exchange. The amount of heat that is dissipated from the body by the different methods varies with the

Fig. 15.3 does not show the effects of humidity, since humidity in the study was moderate. If one includes a high humidity level with high temperature, evaporation also is limited, so the body cannot lose the amount of heat it should lose. The result is extreme discomfort, and reduction in the physical ability to perform tasks.

Fig. 15.4 shows the approximate upper limits of tolerance for heat loss by evaporation for three conditions: resting with air movement, resting without air movement, and working. Combinations of temperature and humidity to the right of a line would preclude evaporative heat loss.

Effects of Atmospheric Conditions on People. The effect of varying temperature upon employee efficiency has not been accurately determined for all types of work, partly because the temperature of the atmosphere, or what is ordinarily known as dry-bulb temperature,

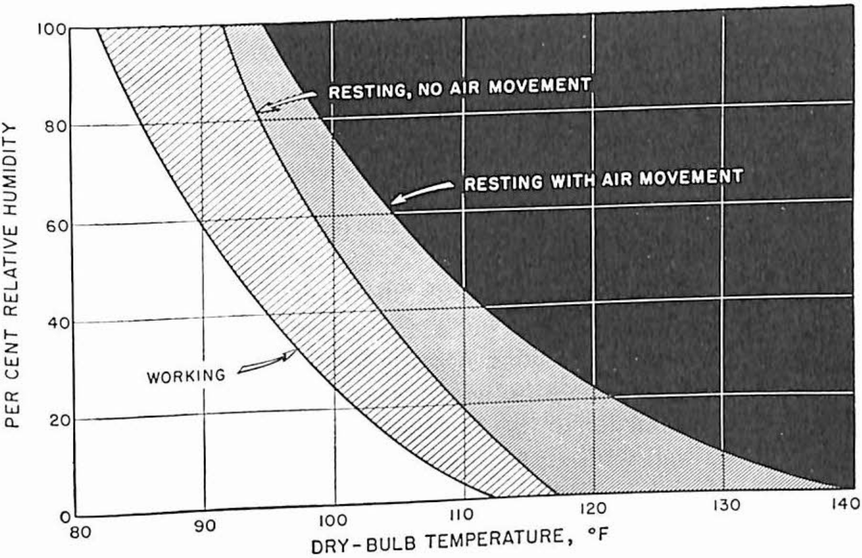


FIG. 15.4. Approximate upper limits for heat loss by evaporation. Each line shows the limit, in terms of temperature and relative humidity, that typically prevents evaporative heat loss for the three conditions. Data are for resting or working nude subjects; for clothed subjects, tolerance limits would be noticeably lower. (Adapted from C. E. A. Winslow et al., "Physiological Influence of Atmospheric Humidity," Second Report of the ASHVE Technical Advisory Committee on Physiological Reactions, Transactions of the ASHVE, 1942, 48, 317–326.)

atmospheric conditions. With high humidity, for example, less heat can be lost by evaporation than with low humidity. Some indication of the relative importance of different methods of heat loss, in relation to atmospheric conditions, is reflected in examples given by Winslow and Herrington.[25] These are shown in graphic form in Fig. 15.3. For each of five combinations of *air temperature* and *wall temperature* there is shown the *per cent* of the total heat loss that occurs by each of three methods: evaporation, radiation, and convection.

With high wall temperatures (such as B, D, and E, with wall temperatures of 120, 126, and 98) the loss of heat by radiation is not possible; the body would be *gaining* heat by this process. With air temperatures rather low (as in A, B, and C) the important role of convection is seen in the figure. With high air temperatures (D and E) the heat-loss burden is thrown almost exclusively on the evaporation process.

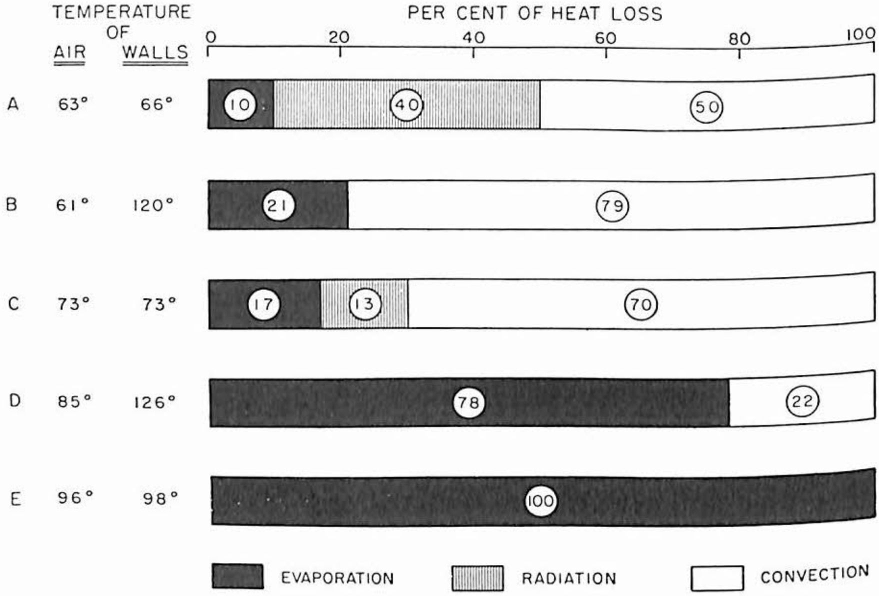


FIG. 15.3. Per cent of heat loss to environment by different methods under different combinations of air and wall temperature. (Adapted from Winslow and Herrington, op. cit., p. 55.)

---

[25] C. E. A. Winslow and L. P. Herrington, *Temperature and Human Life* (Princeton, N.J.: Princeton University Press, 1949).

seldom varies alone. Almost always when atmospheric temperature varies, other conditions, such as humidity and the amount of radiated heat, do not remain constant. In spite of the difficulty of controlling conditions, however, a few studies have shown that temperature is related to employee efficiency. For example, Vernon, Bedford, and Warner[26] found that when temperature was below 70 degrees, only 3 per cent of the time was lost as a result of employee sickness; when temperatures were from 70 to 79.9 degrees, 4.5 per cent of the time was lost; and when temperatures were 80 degrees or more, 4.9 per cent of time was lost as a result of sickness. These findings indicate that the optimal temperature for the plants studied was below 70° F. The optimum, however, presumably varies with the type of activity. Baetjer[27] in summarizing various studies gives the following estimates of desirable conditions for the types of activities in question:

Activity	*Desirable dry-bulb temperature, F°*
Light sedentary work: winter	68–73
Light sedentary work: summer	75–80
Moderate hard work.	65
Strenuous work	60

In these ranges of temperatures, relative humidity is not a particularly important variable.

In giving any such estimates of "desirable" or "optimum" conditions, however, it should be borne in mind that people adapt rather well to a variety of conditions. A study by the New York State Commission on Ventilation[28] indicated, for example, people preferred temperatures from 68° to 75°, but the subjects could perform mental work just as well with an 86° temperature in combination with 80 per cent relative humidity; under these conditions, however, they felt less inclined to work. When performing physical work, however, temperatures of 75° and 86° caused performance to drop considerably, as compared with performance at 68°.

---

[26] H. M. Vernon, T. Bedford, and C. G. Warner, "A Study of Absenteeism in a Group of Ten Collieries," *Industrial Fatigue Research Board Report No. 51* (London, 1928).

[27] Anna M. Baetjer, "Light, Temperature, Humidity," *Industrial Medicine*, 1944, 13, 111–112.

[28] *Ventilation*, Report of the New York State Commission on Ventilation (New York; E. P. Dutton Co., 1923).

The four atmospheric variables that are particularly pertinent to comfort and performance in everyday living are temperature, humidity, radiation (the temperature of surrounding objects such as walls and ceilings), and air flow. The first three have been discussed above in one context or another. Air flow is most important in warm or moderately hot environments, since it exposes the body to changing air, thereby speeding up convection and evaporation heat losses. At temperatures up toward 100°, and above, however, its effect is markedly reduced, and, in fact, in combination with high humidities (which limit evaporation) air flow may even reduce heat loss.[29] In warm environments below such extremes the effect of air flow is cooling, as we all know from sitting in front of an electric fan on a warm day. The increased heat loss from air flow under such circumstances can have a beneficial effect on work performance as indicated by the results of a study in a textile mill by Wyatt, Fraser, and Stock.[30] In their experiment, suitably placed fans were operated on alternate working days over a period of six weeks during the summer. The results indicated that when air was circulated by the fans, production by the employees was significantly higher than when the fans were not in use.

Comfort Zones. Various studies have been made to determine temperature "comfort zones." There are some variations in the temperatures at which people feel comfortable, depending on the season, the age and sex of the people, and their geographical location. Generally speaking, however, the summer comfort ranges are from about 69° to about 73° effective temperature.[31] (Effective temperature is a scale that takes into account both temperature and humidity.) The winter comfort zone is from about 65° to about 70° effective temperature. These effective temperatures can be translated into everyday experiences by considering the given temperatures to be accompanied by 100 per cent relative humidity.

---

[29] *Heating, Ventilating, Air Conditioning Guide,* American Society of Heating and Air Conditioning Engineers, 1957, Chapter 6.

[30] S. Wyatt, F. A. Fraser, and F. G. L. Stock, "Fan Ventilation in a Humid Weaving Shed." *Industrial Fatigue Research Board Report No. 37* (London, 1926).

[31] *Heating, Ventilating, Air Conditioning Guide, op. cit.* Chapter 6.

NOISE

A great deal has been written concerning the various effects of noise in industry, especially on people's work performance and hearing. Before considering these effects, however, we should mention briefly the measurement of sound.

There are two primary characteristics of sound, namely, *frequency* and *intensity*. The frequency of sound is based on the number of vibrations per second of the noise source. Middle C on the piano, for example, has a frequency of 256 cycles per second (cps); each octave has twice the frequency of the octave below it. Most sounds in our environment, however, do not consist of pure tones, but rather consist of combinations of frequencies. Noise has sometimes been characterized as "unwanted" sound.

Intensity is measured by the decibel (db). Fig. 15.5 illustrates the intensities, in decibels, of various noise sources and environments.

Effects of Noise on Work Performance. Various experiments have indicated that work performance improves with noise reduction. A study by Scheidt,[32] for example, suggested that reducing noise levels by noise-proofing workrooms or by putting employees in individual booths can improve the quality of their work. In another study by Weston and Adams[33] the effective noise in a textile mill was reduced from 96 to 87 decibels by the use of earplugs. Considering only the time during which employees were performing nonmachine controlled tasks, there was a 12 per cent increase in employee performance.

The evidence, however, is by no means consistent or conclusive. Stevens and others,[34] for example, studied the effects of simulated

---

[32] V. P. Scheidt, "The Effect of Various Modifications on the Worker," *Kalenda*, 1937, **16**, 3–4.

[33] H. C. Weston and S. Adams, "Two Studies in Psychological Effects of Noise. Part II. The Effect of Noise on Performance," *Industrial Health Research Board Report No. 65* (London, 1932).

[34] S. S. Stevens, *The Effects of Noise and Vibration on Psycho-Motor Efficiency*. Psycho-acoustic Laboratory, Harvard University, March 31, 1941, OSRD Report No. 32.

S. S. Stevens and others, *Part I. The Effects of Noise on Psychomotor Efficiency. Part II. Noise Reduction in Aircraft as Related to Communication, Annoyance, and Aural Injury*. Psychological Laboratory, Harvard University, December 1, 1941, OSRD Report No. 274.

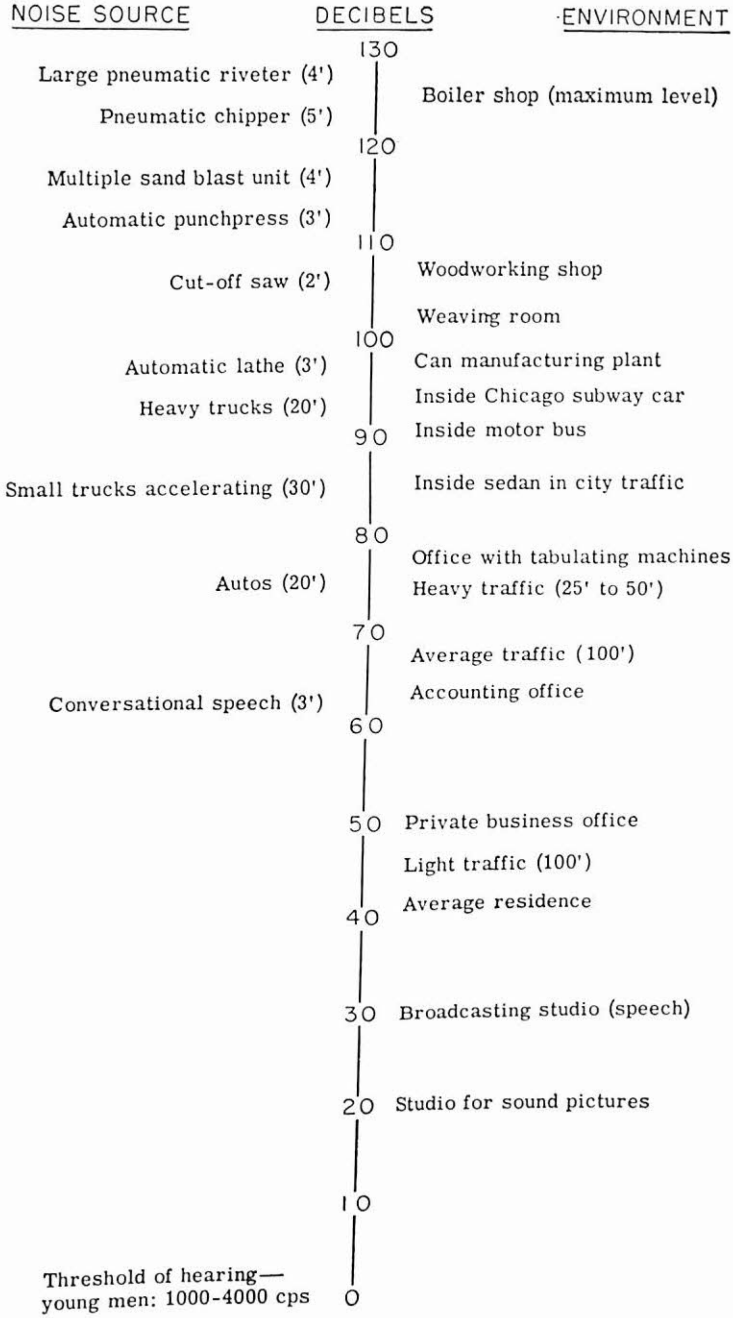


FIG. 15.5. Typical over-all sound levels for various noise sources and environments. (Adapted from A. P. G. Peterson and L. L. Beranek, Handbook of Noise Measurement, General Radio Company, Cambridge, Mass., 1954, p. 2.)

aircraft noise on a group of subjects who were in the noise environment for seven hours a day for four weeks. They found no appreciable change on various motor coordination tests, in reaction time, or on perceptual and mental tests.

Studies such as those by Stevens point up the fact that people have a great deal of adaptability to noise conditions. Such adaptability presumably can serve to maintain work performance at a level comparable to that in more quiet conditions. Thus, the effects of noise on work performance are not unequivocal. Although it has not been possible consistently to demonstrate that high noise levels affect work performance, however, there is at least a strong finger of suspicion pointing in that direction.

Effects of Noise on Hearing. The effects of noise on hearing are much less controversial than are the effects of noise on work performance. Numerous studies and surveys have confirmed the fact that long exposure to high levels of noise brings about some degree of hearing loss.

Typical of various studies of hearing loss is one reported by Glorig and Wheeler[35] of men working in a plant with a noise level of about 105 decibels. Fig. 15.6 shows the average hearing loss of men with various years of exposure to the environment. Each line shows, for the group in question, the average *minimum* intensities of sounds of the various frequencies that could just barely be heard. Fig. 15.6 shows that the greater the length of exposure, the greater the degree of hearing loss, especially in the 4,000 to 8,000 cps range. It should be pointed out that these hearing losses were found to be *greater* than would typically be expected on the basis of age alone. The presumption, then, is that the hearing loss was caused by the plant noise level.

Although precise information is lacking regarding the specific hearing loss that would be brought about by any given noise condition, there is general support for saying that the amount of hearing loss is affected both by intensity of the noise and by the duration of the exposure to the noise.[36] The opinions of experts do not yet support the notion of a neat dividing line between "safe" and "harmful" noise

---

[35] A. Glorig and D. E. Wheeler, "An Introduction to the Industrial Noise Problem," *Illinois Medical Journal*, January 1955, **107,** No. 1.

[36] *The Relations of Hearing Loss to Noise Exposure* (New York, American Standards Association, Inc., 1954), p. 30.

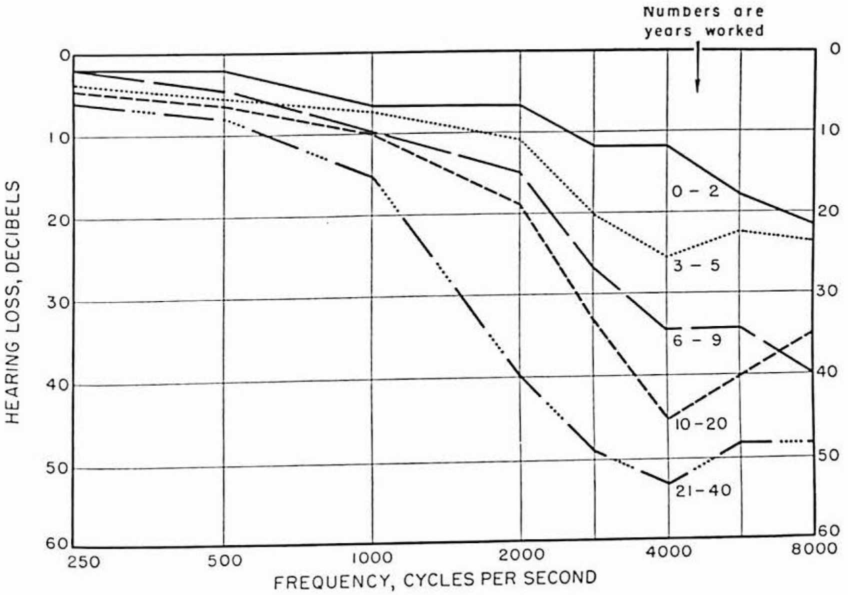


FIG. 15.6. **Average hearing loss of men working, for years shown, in noise environment of about 105 decibels. (From Glorig and Wheeler, op. cit.)**

conditions, as pointed out by Sterner.[37] In a survey he solicited the opinions of experts regarding the "safe" and "harmful" limits, and he received widely differing opinions. The median "safe" estimate was 85 decibels (meaning that, of those surveyed, 50 per cent gave lower values as "safe," and 50 per cent gave higher values); however, there were very few estimates that were higher than 90 decibels, or lower than 80 decibels.

Evidence from one study[38] lends factual support to the judgments of experts that noise levels of about 80 decibels generally are "safe" (they are not accompanied by appreciable hearing losses). Noise levels of 88 decibels, however, were accompanied by some degree of hearing loss, especially at the 4,000 cps level; and at this level even 83 decibels brought about noticeable loss, increasing with the exposure period.

---

[37] J. H. Sterner, "Standards of Noise Tolerance," in *Noise, Lectures Presented at the Inservice Training Course on the Acoustical Spectrum,* University of Michigan School of Public Health and Institute of Industrial Health, February 5–8, 1952, pp. 161–166, University of Michigan Press, Ann Arbor.

[38] *The Relations of Hearing Loss to Noise Exposure, op. cit.,* pp. 31, 32, 33.

Noise Control. The objective of a noise control program in industrial and business organizations is not the elimination of noise, but rather its reduction to reasonable levels, consistent with the type of activity being carried out, and with employee welfare and performance. For extremely noisy circumstances, noise control preferably should be directed *at least* to the reduction of noise to levels that will not cause appreciable hearing loss. Granting that the "safe" level is not highly precise, nonetheless we have seen that noise in the 80 to 90 db level is not generally considered to be "safe."

Noise control can be achieved by various methods, frequently used in combination. These include reduction at the source (such as by proper machine design, proper maintenance, lubrication, mounting equipment on rubber); enclosing the noise; use of baffles; use of sound-proofing materials (walls, ceilings, floors, and so on); and the use of ear plugs.

MUSIC

A number of surveys have been conducted and reported to show how employees feel toward music while they work. Prominent among these has been the work of Kerr[39] and Kirkpatrick.[40] In one of the investigations it was found that approximately two-thirds of the employees stated that they preferred to work with music, whereas nearly all of the remaining one-third were indifferent. Practically none reported that they preferred to work without music. Of those who preferred to work with music, only 14 per cent preferred to have music played all the time. Spears[41] conducted a survey of 262 plants and found that 111 were using some kind of music program. Of these 111, only 9 per cent were dissatisfied with the program. Very few, however, had attempted specific evaluations of the effect of the music. Beckett[42] concludes from a survey of 78 plants that music is universally liked by the workers. A study by Kerr[43] based on a questionnaire survey of 666 employees in a plant of the Radio Corporation of America, showed that the average worker likes between six and seven hours of music on an eight-hour shift.

---

[39] W. A. Kerr, "Psychological Effects of Music as Reported by 162 Defense Trainees," and "Factor Analysis of 229 Electrical Workers' Beliefs in the Effects of Music," *Psychological Record*, 1942, **5**, 205–212, 213–221.

[40] F. H. Kirkpatrick, "Music in Industry," *Journal of Applied Psychology*, 1943, **27**, 258–274.

[41] E. M. Spears, "Music in Industry," *National Industrial Conference Board, No. 78* (1947), 1–32.

[42] W. Beckett, *Music in War Plants* (Washington, D.C., War Production Board, 1943).

Music may, however, be liked by employees even when it has no effect upon the quantity or quality of their work, or even, conceivably, when it has an unfavorable effect upon their efficiency. A number of investigations have therefore been conducted to determine the effect of music upon worker efficiency. Wyatt and Langdon[44] found the production of a small group of women on light repetitive work to be increased from 6 to 11 per cent by varying amounts, distributions, and types of music. Humes[45] concluded from another study that music tends to lower the scrappage rate. Kerr[46] in a series of studies, found that music increases production up to 6 per cent on operations paid on an incentive plan, and up to nearly 12 per cent on hourly-paid jobs. A study by Smith[47] has further demonstrated the favorable effect of music upon production when the work in question is repetitive. No significant differences were found between accident experience in plants with industrial music and in those without it.

HOURS OF WORK

During the past fifty years the trend toward a reduced number of hours in the working week as well as a shortened working day has been fairly constant. This change has been due partly to the feeling of labor that life should not be "all work and no play," but it also has been due in no small measure to an increased recognition by management that employees do not reach their greatest efficiency when the number of working hours is excessive. Many experiments have been conducted to determine the relation between the hours of work and employee production. The reduction in hours of work has been largely justified by these experiments. Miles and Angles[48] report that when the hours of work in a plant manufacturing boxes were reduced from 48 to 36 per week, the average hourly output increased from 793.5 to 834.0 units of production. In another experiment[49] dealing with women engaged in the work of turning fuse bodies, a 68 per cent increase of hourly output and an increase of 16 per cent in total production followed a reduction from sixty-six to forty-eight and six-tenths in the number of hours worked per week. In another study[50] it was found that an increase in working time above sixty hours per week was accompanied by increase in time lost owing to sickness, injury, and absence without permission. The increase in working time was also accompanied by a decrease in both hourly and weekly output and an increase in labor turnover. Murray[51] reports that a reduction of the working hours to less than forty-eight per week was followed by increases in both hourly and weekly production as well as by decreases in absenteeism and turnover. The general conclusion suggested by these and similar studies is that output frequently is increased if the number of working hours per week does not much exceed forty.

---

[43] W. A. Kerr, "Worker Attitudes Toward Scheduling of Industrial Music," *Journal of Applied Psychology*, 1946, **30**, 575–578.

[44] S. Wyatt, J. N. Langdon, and F. G. L. Stock, "Fatigue and Boredom in Repetitive Work," *Industrial Health Research Report No. 77* (London, 1937), 30–42.

[45] J. F. Humes, "The Effect of Occupational Music on Scrappage in the Manufacture of Radio Tubes," *Journal of Applied Psychology*, 1942, **25**, 573–587.

[46] W. A. Kerr, "Experiments on the Effects of Music on Factory Production," *Applied Psychology Monographs, No. 5* (1945), 1–40.

[47] H. C. Smith, "Music in Relation to Employee Attitudes, Piecework Production and Industrial Accidents," *Applied Psychology Monographs, No. 14* (1947), 59 pp.

Of course, during a period of emergency production, it is not always possible to achieve maximum efficiency by means of a short working week. At such times it is often necessary to get out maximum production even though by so doing maximum efficiency, from the long-time viewpoint, may be sacrificed to some extent. But even during a period of maximum production the question may be raised whether the real gain can be achieved by excessive lengthening of work hours. Ivy[52] has suggested that from the physiological viewpoint we will sacrifice rather than gain, even in terms of short-term production, by increasing the hours of work beyond fifty-six hours per week for men and forty-eight hours per week for women. Miller[53] found that even under wartime conditions, hours for women cannot be increased to more than around forty-eight hours per week without lowering productive efficiency.

---

[48] G. H. Miles and A. Angles, "The Influence of Short Time on Speed of Production, II," *Journal of the National Institute of Industrial Psychology*, 1925, **2**, 300–302.

[49] C. S. Myers, *Mind and Work* (London, Putnam & Company, 1921).

[50] Anon., "Hours of Work, Lost Time, and Labour Wastage," *Industrial Health Research Board, Emergency Report No. 2* (London, 1942).

[51] H. M. L. Murray, "Bases of Worker Efficiency," *Personnel Journal*, 1942, **21**, 131–145.

[52] A. C. Ivy, "The Physiology of Work" (Fourth Annual Congress on Industrial Health, Chicago, January 1942).

REST PAUSES

A great many investigations on the effect of rest pauses have been conducted. These studies have considered employees on both heavy and light work, and employees paid under a straight hourly rate as well as employees paid under a wage incentive plan. In this field the work of Vernon[54] and his associates is outstanding. Their studies have shown that rest pauses usually result in increased production in spite of the fact that the pauses result in some reduction of the actual working time. Typical curves showing the production for a group of employees before and after the introduction of rest periods are shown in Fig. 15.7. These curves are based upon an investigation of Farmer and Bevington.[55]

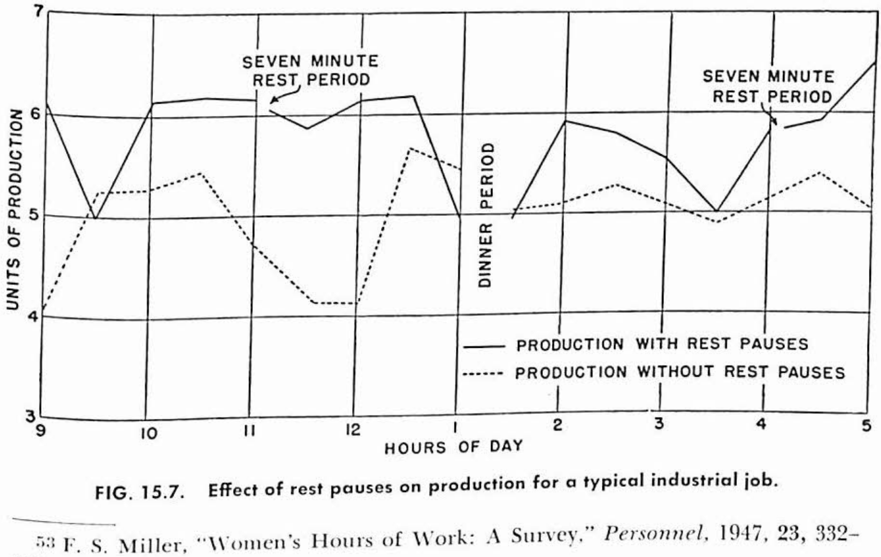


FIG. 15.7. **Effect of rest pauses on production for a typical industrial job.**

---

[53] F. S. Miller, "Women's Hours of Work: A Survey," *Personnel*, 1947, **23**, 332–338.

[54] H. M. Vernon, *Industrial Fatigue and Efficiency* (New York, E. P. Dutton & Co., Inc., 1921).

[55] E. Farmer and S. M. Bevington, "An Experiment in the Introduction of Rest-Pauses," *Journal of the National Institute of Industrial Psychology*, 1922, **1**, 89–92.

In another study, Miles and Skilbeck[56] found that two fifteen-minute change-of-work periods resulted in a 14.2 per cent efficiency increase. Rest periods should be introduced at those times when employees are experiencing the greatest fatigue. An investigation similar to one reported by Griffith, Kerr, Mayo, and Topal[57] is often helpful in identifying the proper times. These investigators found that 379 employees in one plant, surveyed by means of an anonymous questionnaire, rather generally reported greatest feelings of "fatigue" during the fourth and eighth hours of the eight-hour shift. The results were remarkably similar for the manual, office, and supervisory employees covered in this study.

It is impractical to generalize about how long rest periods should be for any given plant, or how frequently they should be allowed during the day. Both the duration and the number of rest periods necessarily vary with such factors as the nature of the work, the effect of a temporary stoppage upon the flow of production, and the availability of suitable rest rooms or other places in which the employees can relax. The duration of rest periods that have proved beneficial has varied from two minutes to fifteen minutes under different conditions, and the number of rest periods during the day has varied from one to five or even more. An excellent summary of representative studies in this field has been given by Viteles[58] and will not be repeated here. Anyone who wishes to consider the introduction of rest pauses for an industrial job should make a careful study of the literature to find whether there are any experiments reported in which rest pauses have been introduced on a somewhat similar job. The results of any such study then might be of use in determining the frequency and duration of rest pauses for the situation at hand. Generally speaking, most investigations on this subject have indicated that some form of rest pause or pauses has been found to be worth while from the point of view both of employee comfort and of employee performance; typically, the time out for rest pauses is more than made up in terms of improved employee performance.

---

[56] G. H. Miles and O. Skilbeck, "An Experiment on Change of Work," *Occupational Psychology*, 1944, **18**, 192–195.

[57] J. W. Griffith, W. A. Kerr, T. B. Mayo, Jr., and J. R. Topal, "Changes in Subjective Fatigue and Readiness for Work During the Eight-hour Shift," *Journal of Applied Psychology*, 1950, **34**, 163–166.

[58] M. S. Viteles, *Industrial Psychology* (New York, W. W. Norton & Company, Inc., 1932), 470–482.

In considering the introduction of systematic rest pauses, it should not be assumed that if no rest pauses are formally allowed the employees will not rest. Several investigations have shown that if formal rest pauses are not allowed, unauthorized rests will be taken by the employees to suit their own convenience.[59] In some cases, such unauthorized rests are as effective in increasing production and allaying fatigue as rest pauses that are officially and formally allowed by management. But in other cases, employees who decide upon their own rest periods may do so in an ineffective manner, or may inconvenience the production schedule by resting at the wrong time. McGehee and Owen[60] found that the introduction of two short rest pauses during the day reduced unauthorized rest pauses and increased the speed of work. In general, it is a more satisfactory procedure for management to decide upon the proper amount and distribution of rest periods and then formally to allow those rests. Just as a person left to learn a job himself may, and often does, adopt an inefficient and cumbersome way of performing the job, so an employee left to decide upon optimal rest pauses for himself will not always adopt rest pauses that are of maximum value either to himself or to departmental production.

In connection with rest pauses, some companies provide food, either free or at nominal cost, as a means of providing employees with a nutritional "pick-up." Such nourishment may be in the form of snacks, sandwiches, milk, coffee, and in some cases even extra meals. Haggard,[61] for example, reports that when employees in one plant were given two extra meals, output rose by 10 per cent, and the workers reported feeling less tired in spite of their greater output.

---

[59] R. B. Hersey, "Rests—Authorized and Unauthorized," *Journal of Personnel Research*, 1925, **4**, 37–45.

[60] W. McGehee and E. B. Owen, "Authorized and Unauthorized Rest Pauses in Clerical Work," *Journal of Applied Psychology*, 1940, **24**, 605–614.

[61] W. W. Haggard, "Work and Fatigue," *Mechanical Engineering*, 1936, **58**, 298–301.

in some respects human engineering is not a particularly appropriate name, it is the one that seems to be most widely used, and so we will use it here.

MAN-MACHINE SYSTEMS

In a human engineering context, human beings are considered as being part of a *man-machine system*. In such a system a man performs essentially three types of processes, as follows: (1) information-receiving, (2) decision-making, and (3) action-taking. Consider, for example, an electric power-station operator. He receives information from the instruments of his control panel. On the basis of this information he decides what to do, and he takes an action such as closing one switch and opening another. The basic relationship of men to their work situations is shown in Fig. 16.1. Although work situations naturally vary a great deal, these three basic types of functions are performed on every job.

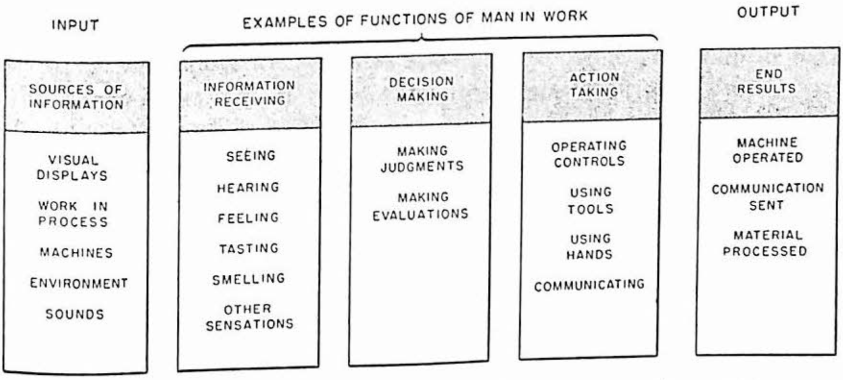


FIG. 16.1. **Schematic diagram of functions of man in man-machine system.**

The objectives of human engineering, in general, are those of so designing the equipment, work space, and environment as to facilitate the information-receiving, decision-making, and action-taking functions of men in the system in order to optimize the "output" of the total system.

The range of studies dealing with different aspects of human engineering is broad and varied. Our purpose here will be to give a few

16

HUMAN ENGINEERING AND WORK METHODS

Throughout history men have endeavored to develop ways and means of performing work more effectively and more easily. Such efforts have been directed especially toward the improvement of the design of equipment and of methods of work. In connection with equipment design, for our purposes here we should differentiate between strictly engineering considerations and design features relating to human use. Our interest with respect to design will be in terms of human use.

Improvements throughout history in the design of equipment for human use and in work methods have come about primarily through man's experience. He has found, for example, that the hoe, as it has evolved to its present design, is a reasonably satisfactory tool for its purpose. However, although experience has been very useful in the development of equipment and work methods, it has its limitations. This is particularly the case with many new types of equipment, with which there has been very little opportunity to benefit by "experience." In large part the limitations of experience can be supplemented through systematic research. In recent years, especially during and since World War II, there has been a great increase in the application of research findings to problems of designing machines for human use. Much of this research has been stimulated by the need for designing new, complex military equipments in such a way that they can be operated effectively by military personnel.

The term *human engineering* has come into use in recent years to apply to the design of equipments, work areas, and even work environments in terms of human abilities and limitations. Although illustrative examples of human engineering studies,[1] not to cover the whole subject. Such examples may illustrate at least a few of the possible applications to engineering design of human engineering research.

INFORMATION-RECEIVING PROCESSES

The sensory organs of the body are the avenues through which an individual receives information regarding the world about him, including the information that is available to him in performing his job. We commonly think of there being five senses, namely seeing, hearing, touch, taste, and smell. There are, however, certain other senses, such as sensations of heat and cold, of body movement, of body posture, and of position of body members (and probably some others not yet identified). Seeing and hearing are the senses used most frequently in work situations, although some of the other senses are important on certain kinds of jobs.

In designing the equipment for a job there are, in general, two major considerations in determining what sense or senses to use in the transmission of information to the worker. In the first place, in some circumstances the appropriate sense is, for practical purposes, suggested by the nature of the source of information or by the circumstances under which it is to be received. The fact of a telephone call coming in, for example, can most appropriately be made known to the household by hearing, because of the attention-getting power of sound throughout a house. We might also cite the instance of an electric meter in the home. To convey information about electric-power consumption in the home, one would certainly not want a noise stimulus; thus, a meter provides a record that can be seen visually. These two examples are clear-cut; but there are many other situations in which the appropriate sense avenue is not nearly as straightforward. For example, although information about the speed of the family car typically has been presented visually, one could argue for the advantage of having a sound signal to indicate an appropriate maximum speed.

---

[1] Readers interested in further information on human engineering are referred to other sources, such as:

A. Chapanis, W. R. Garner, and C. T. Morgan, *Applied Experimental Psychology* (New York, John Wiley & Sons, Inc., 1949).

E. J. McCormick, *Human Engineering* (New York, McGraw-Hill Book Co., Inc., 1957).

W. E. Woodson, *Human Engineering Guide for Equipment Designers* (Berkeley. University of California Press, 1954).

Handbook of Human Engineering Data, (2nd ed., rev.; Tufts College Institute for Applied Experimental Psychology, 1952) (available through Tufts College Bookstore, Medford, Mass.).

In the second place, where either of two (or more) sense avenues could be used in the transmission of information, the determination of the method to be used may be based on consideration of the relative demands already imposed on the various sensory avenues. Where, for example, the visual sense is already burdened, perhaps the auditory or the tactual senses might be brought into use to relieve the visual sense. This would be particularly desirable in complex operations where a great deal of information has to be received and integrated, such as in the operation of aircraft and some types of military equipment.

Visual Displays. Because of the greater emphasis in human engineering research to date on the visual method of information-receiving, our examples of the information-receiving phase of work will deal with visual displays.

The term *visual display* applies to any device that is used to present information visually to human beings. Examples include dials and other visual instruments, signal lights, signs, and so forth. A great deal of the research in human engineering to date has been directed toward the design of visual displays that convey the information most adequately to human beings.

Instrument design. An example of such research is a study reported by Sleight[2] in which instruments of five different shapes were used. The purpose was that of determining which type could be read most accurately. The five designs used are shown in Fig. 16.2. The features of the five designs are identical (numerals, size of graduation, distance between graduations, and design of pointers).

Sixty male subjects were used, each being presented with 17 pointer settings on each of the five designs, each setting being shown for .12 second. Errors in reading were recorded; they are summarized in Fig. 16.3. This shows the clear superiority, in reading accuracy, of the open-window design. The vertical and horizontal designs were particularly inaccurate.

Arrangement of dials for check reading. Sometimes many dials are

---

[2] R. B. Sleight, "The Effect of Instrument Dial Shape on Legibility," *Journal of Applied Psychology,* 1948, **32,** 170–188.

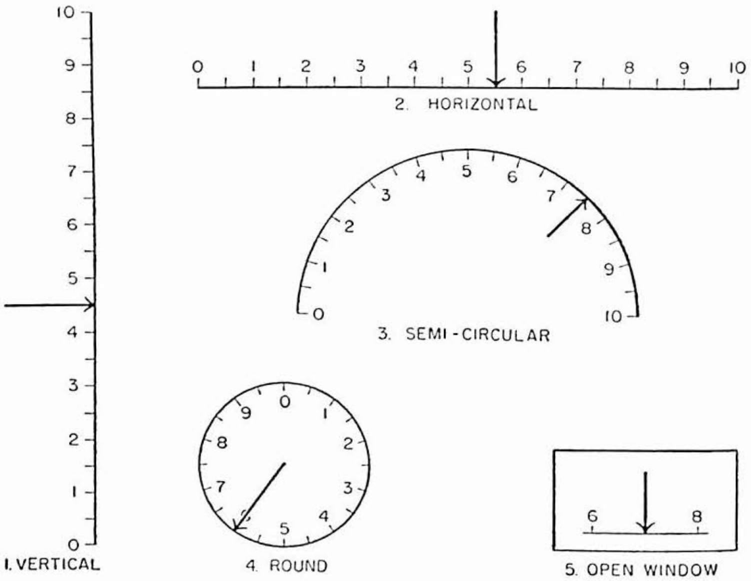


FIG. 16.2. **Illustrations of dials used in study on legibility. (From Sleight, op. cit.)**

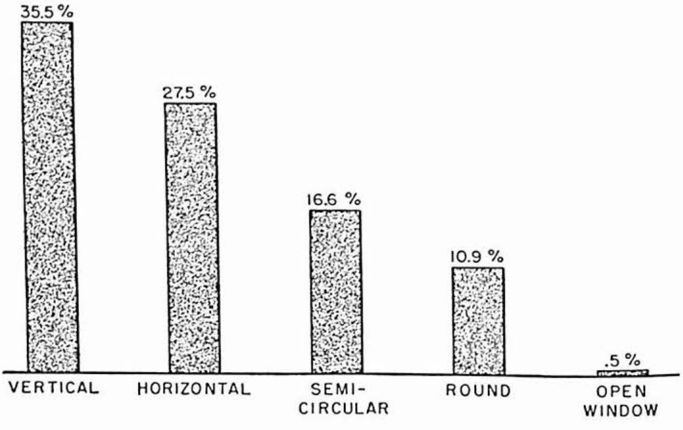


FIG. 16.3. **Per cent of errors in reading dials of five different designs. (From Sleight, op. cit.)**

arranged together into a bank of dials. In some such instances the dials are used for "check reading," which is simply the determination as to whether each dial is at a "normal" condition, or not.

In a study by Senders[3] a comparison was made of the time required for reading groups of dials when the "normal" position in all cases was at the same relative location (namely, at the 9 o'clock position), as opposed to reading groups of dials when the "normal" position was different for each dial. She found that 32 "aligned" dials could be read in the same amount of time (0.5 second) as 4 "nonaligned" dials.

In another study, Johnsgard[4] in turn, investigated the speed of checking nonaligned pointers of groups of dials arranged in four different ways, as shown in Fig. 16.4. Each subject was presented with 18 sheets with dials arranged in each of the four patterns, making a total of 72 sheets. Some of the sheets had pointer settings that were *not* in the "normal" position, and the subject was to identify such out-of-normal dials. Each subject saw a sheet for half a second.

The mean accuracy scores for the four patterns are shown by the bars underneath each of the four patterns in Fig. 16.4. This shows that the subgroup patterns were read less accurately than the other three patterns. Although the differences among the other three are small, pattern 3 (one of the "symmetry" patterns) was slightly superior to the others.

DECISION-MAKING PROCESSES

The distinction between the process of receiving information (which is, initially, a sensory process) and that of making decisions based on such information is difficult to draw with precision. This is because the simple sensory process of receiving a stimulus through a sense organ is inextricably intertwined with *perception*. Perception, as a psychological function, involves the attachment of *meaning* to that which is sensed. Let us take, for example, two people looking at the pressure-gauge of a steam boiler, both of whom are equally able to "see" the reading. One person knows nothing about boilers, so to him the information is meaningless. The other one, who knows about such things, sees that the pressure is very near what he knows to be the danger point. The "perceptions" of these two people of the same,

---

[3] Virginia L. Senders, *The Effect of Number of Dials on Qualitative Reading of a Multiple Dial Panel*, USAF, Wright Air Development Center, Technical Report 52–182.

[4] J. W. Johnsgard, "Check-Reading as a Function of Pointer Symmetry and and Uniform Alignment," *Journal of Applied Psychology*, 1953, **37**, 407–411.

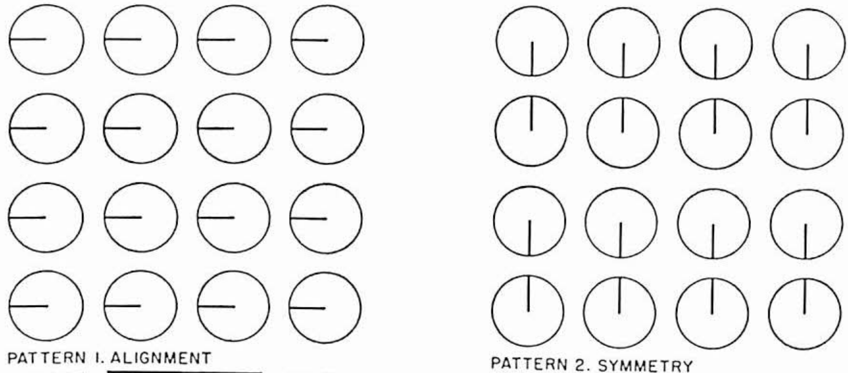


FIG. 16.4. Patterns of dials used in check-reading experiment, showing mean accuracy scores in identifying out-of-normal positions. (Adapted from Johnsgard, op. cit.)

identical instrument, would be very different. The decision, on the part of the second person, to reduce the pressure, would be an almost automatic one.

The making of decisions in a work situation probably depends, essentially, upon four factors, as follows: (1) the potentialities of the individual to make the decisions in question (his aptitudes, and the like); (2) his training; (3) the nature of the decision to be made; and (4) the manner in which the information is presented. The first two of these deal primarily with the individual (his aptitudes and training). Engineers, psychologists, and others who are concerned with human engineering are particularly interested in the latter two factors, the nature of the decisions to be made and the manner in which the information is presented to the worker.

Nature of the Decisions to Be Made. The nature of decisions made on various jobs ranges from those which are extremely simple (such as deciding to insert a cotter pin in the hole provided on the axle of a tricycle), to those decisions which, let us say, the President might make in establishing foreign policy. The "decision" to place the cotter pin in the hole is, in one sense, hardly a decision at all. It is a prescribed action to take place under a prescribed circumstance.

The point we are leading up to is that, generally speaking, there is an advantage in the operation of man-machine systems in providing for *standardized* decisions to be made for specified conditions (as indicated by the "information" presented to the operator).

Further, the decision-making process is more rapid when the *number* of possible choices of action (even when each is standardized) is limited. In this connection Fitts[5] has summarized a number of pertinent studies, including one by Hick.[6] In this study the subjects were asked to respond to appropriate signals by moving fingers that corresponded to the signals. The number of signals (and corresponding number of fingers) varied from 1 to 10 in different parts of the experiment. Thus, if six signals were being used, the subject, upon seeing a particular signal, would have to move the finger that corresponded with that signal. The time required for making such choices is given in Fig. 16.5. These times are usually referred to as *reaction time*, but we can think of this as "decision" time. We can see that the decision time varies with the number of choices from which a selective choice is to be made. The decision time, incidentally, increased logarithmically with increases in number of alternate choices.

This study probably represents an oversimplification of the general principle in question, but even so it suggests that, where the decisions (choices of responses) are to be made on the basis of presented information, there is a time advantage (and reduced likeli-

---

[5] P. M. Fitts, "The Influence of Response Coding on Performance in Motor Tasks," in *Current Trends in Information Theory* (Pittsburgh, University of Pittsburgh Press, 1953).

[6] W. E. Hick, "On the Rate of Gain of Information," *Quarterly Journal of Experimental Psychology*, 1952, **4**, 11–26.

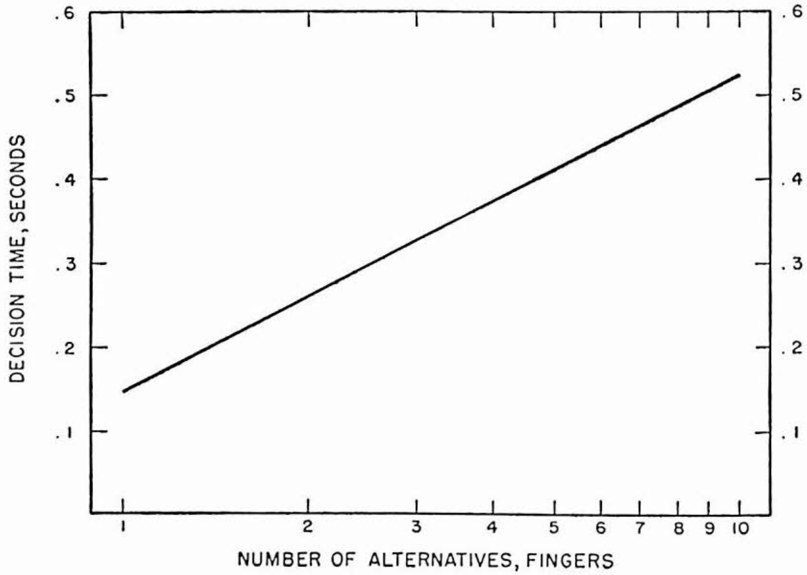


FIG. 16.5. The relationship between decision time and number of possible alternate responses from which to choose in making responses to appropriate signal. (Data from Hick, op. cit., as adapted by Fitts, op. cit.)

hood of error) when the number of alternate choices available is limited.

Manner of Presenting Information. Granting then that there is an advantage, in the design of a man-machine system, in providing for simple, direct decisions where possible rather than complex ones, the question for the designer is how this can be accomplished. In part the answer to this question goes back to the presentation of information to the worker. It would be desirable to present information to an operator in such a way that it is most *effective* in facilitating the decision-making process. It must be stated that much needs to be learned about this. An example, however, may suggest the type of analysis that can contribute some answers to this question.

A study was carried out by Forbes[7] for purposes of studying three auditory signals used in flying. We will not bother to discuss the procedures of the experiment, or the specific results, but rather we will simply summarize the implications. It was found that a combination three-in-one sound signal was better than three separate signals.

---

[7] J. W. Forbes, "Auditory Signals for Instrument Flying," *Journal of the Aeronautical Sciences,* 1946, 255–258.

The combination signal would be varied in three different ways (specifically in steadiness, frequency, and modulation) to provide information of three different types. Such integration apparently contributed to more effective decision-making than having each type of information presented by a separate, independent signal.

The implications of this type of research strongly suggest that the appropriate *integration* of information for workers can facilitate the decision-making process.

ACTION-TAKING PROCESSES

Of the human-engineering research regarding action-taking processes, probably the greatest emphasis to date has dealt with the design and arrangement of control devices or with the related physical motor abilities of people. The examples given in this section, then, will be those associated with control devices and motor abilities.

Blind-Positioning Movements. Frequently on a job the worker makes what is called a "blind-positioning" movement to some control device such as a lever or push button without looking, or, in some circumstances, without being able to look. The accuracy with which blind-positioning movements can be made in different directions is then a potentially important consideration in deciding where the control device should be located.

A study by Fitts[8] dealt with this. The subjects in the experiment were blindfolded, then seated in the enclosure shown in Fig. 16.6. They were given a sharp-pointed marker and were asked to make blind-positioning movements to the various targets around the enclosure. There were three tiers of targets, the middle one being at approximately shoulder level, and the others being about 45° above and below.

Accuracy was determined by "scoring" the position of the hits made with the marker; a bull's eye was scored 0, a mark in the first circle around the bull's eye was scored 1, and marks in the other rings were given higher scores up to the outside ring which was scored 6. Thus, a low score indicates good accuracy. The results are shown in Fig. 16.7. The size of each circle indicates the relative error-scores of the movements to the target at that particular location. The average

---

[8] P. M. Fitts, "A Study of Location Discrimination Ability," in *Psychological Research on Equipment Design*, ed. P. M. Fitts, USAF, Aviation Psychology Program, Research Report No. 19, Superintendent of Documents, 1947.

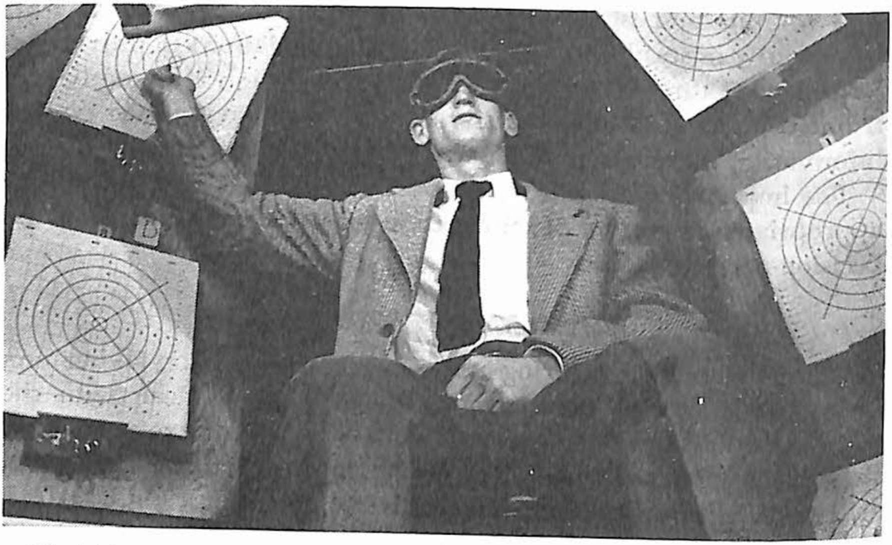


FIG. 16.6. Experimental situation used in study of accuracy of blind-positioning movements carried out by Fitts, op. cit. (Courtesy Psychology Branch, Aero-Medical Laboratory.)

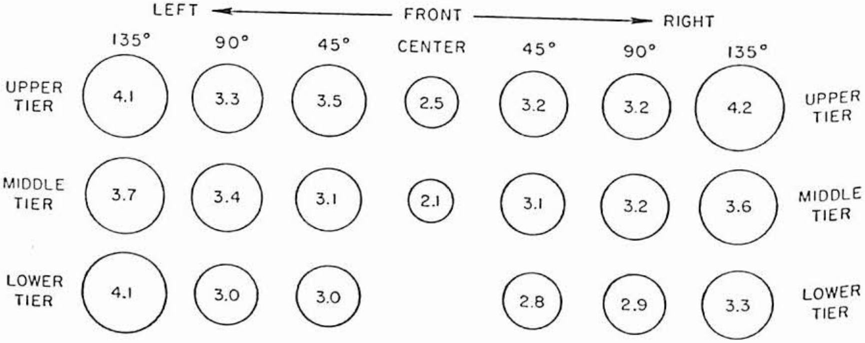


FIG. 16.7. Accuracy of blind-positioning movements in different directions. Each circle represents the location of a target (left to right, and in upper, middle, and lower tiers). The average error-score for each target is given inside the circle, and the size of the circle is proportional to the error-score. (Adapted from Fitts, op. cit. Courtesy Psychology Branch, Aero-Medical Laboratory.)

error-scores are given inside the circles. Fig. 16.7 indicates that accuracy is greatest in the positions straight ahead and is least in the extreme side positions. Considering target level, the accuracy is greatest for the lowest tier and least for the highest tier.

The results of a study such as this suggest that when positioning control devices that are likely to be used without looking, greater accuracy in locating them could be achieved by placing them toward the center (ahead of the worker or slightly to his right or left), and preferably below shoulder height.

Size of Handwheels. Many machines or other equipments have handwheels that serve some function such as moving the carriage of the machine to a particular position. Since the purpose and method of use will in large part determine the suitability of various sizes of handwheels, it is obviously not possible to use the same size of handwheel for all purposes. Some, for example, have to be rotated rapidly for many turns; others have to be used to overcome considerable resistance; still others may have to be used with great precision.

To illustrate the type of study that can be carried out in analyzing handwheels for a particular type of purpose, let us cite a portion of a study by Davis.[9] He experimented with various sizes of handwheels which the subjects were to use for the precise positioning of a pointer. Generally speaking the movement required about one revolution. Another variable that he introduced was the amount of torque in the wheels (0, 40, and 90 inch-pounds). He recorded the time required to make settings with different combinations of size and torque, the results being shown in Fig. 16.8. It can be seen that with no torque there is not much difference in time for making a setting among the various sizes, although the 3-inch and 6-inch handwheels took slightly less time than the larger ones. When using handwheels with substantial torque, however, the size of the handwheel is a very definite factor in the speed of use. Although with a 40 inch-pound torque the 10-inch and the 16-inch handwheels gave equal times, with the 90 inch-pound torque the 16-inch handwheel was distinctly superior to the others.

The results of an analysis of this type would be very helpful to a design engineer in selecting a handwheel that would be most rapidly used (under a given torque condition) for a purpose such as the one used in this study.

Workspace. An important consideration in the performance of manual work activities is that of the physical space within which the work is to be done. An illustration of research carried out in this phase of

---

[9] L. E. Davis, "Human Factors in Design of Manual Machine Controls," *Mechanical Engineering*, October 1949, 71, 811–816.

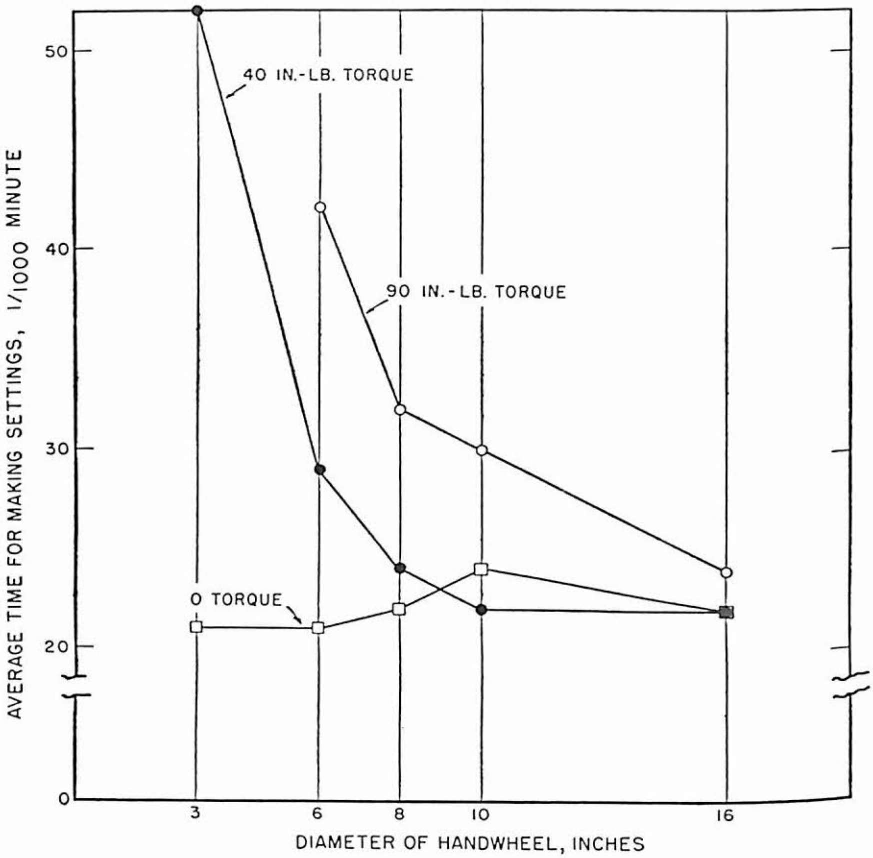


FIG. 16.8. Average times required for making pointer settings with handwheels varying in size and torque. (Adapted from Davis, op. cit.)

human engineering is a recent study by Squires[10] in which he was concerned with the shape of the "normal work area," such as on a work bench or table. The normal work area traditionally has been considered to consist of overlapping semicircles, each one being formed by the rotation of the forearm with the elbow as a pivot.[11] The study by Squires, however, indicates that the normal work area is somewhat more limited than is traditionally thought. This area

---

[10] P. C. Squires, "The Shape of the Normal Work Area," *U.S. Naval Medical Research Laboratory Report No. 275*, 23 July 1956.

[11] R. M. Barnes, *Motion and Time Study* (3rd ed.; New York, John Wiley and Sons, Inc., 1949), Fig. 144, p. 233.

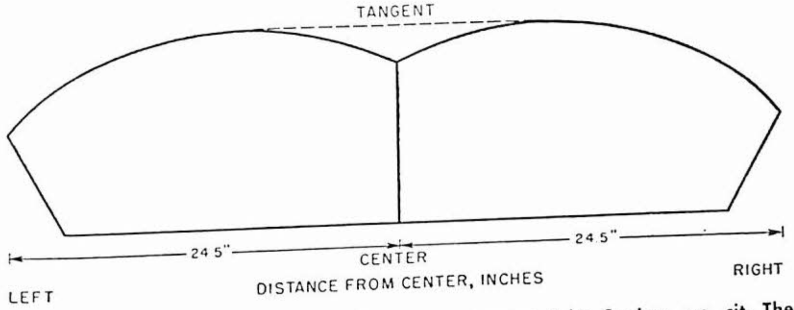


FIG. 16.9. Normal horizontal work area as presentsd by Squires, op. cit. The curves show the areas that fall within convenient reach of the hands when the forearms are rotated around the elbow. The elbow itself tends to move during such rotation.

is shown in Fig. 16.9.[12] It takes into account the fact that the elbow typically moves out and to the side as the forearm is rotated about it. It also shows the effect of limiting outer position of the forearm-hand as it rotates outward.

This shape is suggested by Squires not only as the normal work-surface area of a work station, but also as the contour of a vertical panel with control devices. Such a proposed panel is shown in Fig. 16.10. Here the horizontal surface has been made straight, along the tangent of the two curves shown in Fig. 16.9; and the corresponding vertical surface in Fig. 16.10 has been made correspondingly flat.

METHODS ANALYSIS

The industrial engineers have developed various techniques for analyzing job methods in order to improve them. These techniques include motion and time study. The purposes of motion and time study, as given by Barnes,[13] are as follows:

1. Finding the most economical way of doing the work.
2. Standardizing the methods, materials, tools, and equipment.
3. Accurately determining the time required.
4. Assisting in training the worker in the new method.

---

[12] For the mathematically inclined, it might be added that the curve is a *prolate epicycloid.*

[13] Barnes, *op. cit.*, p. 1.

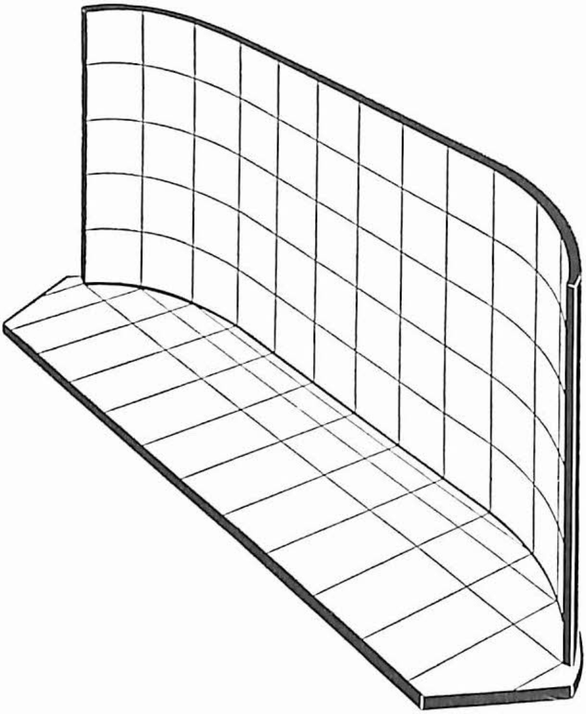


FIG. 16.10. Horizontal and vertical work surface areas as adapted from Fig. 16.9. (Based on study of normal work areas by Squires, op. cit.)

From ordinary observations and the application of simple logic it is apparent that different methods of doing a certain job may require different amounts of time and physical energy. If an individual has a pile of rocks to move from one location to another, for example, he can do the work with less physical exertion by using a wheelbarrow than by carrying the rocks one at a time in his hands.

The methods of motion and time study involve the systematic analysis of work activities (rather than "observations") in order to bring about improvements, which may not be obvious from unsystematic observations. The possibilities of improvement by such methods were dramatically demonstrated early in the century by Gilbreth[14] with the operation of bricklaying. He showed that the work of the average bricklayer can be increased from 120 bricks per hour to 350 bricks per hour by following a more efficient pattern of movement. This early study with its concrete results pointed the way toward an application of what is now known as motion and time study in a wide variety of industrial jobs. Among the possible variations in work method are different types of bodily movement that may be required with different methods of operation. The bodily movement aspect of motion study may be considered a legitimate field for psychological investigation[15] although it is ordinarily considered a branch of industrial engineering. Since instruction and laboratory work in this field are usually given by departments of industrial engineering, a thorough coverage of the topic does not come within the scope of this book. The industrial psychologist should, however, be aware of the significance of the field and should supplement his training in psychology with sufficient work in motion and time study to make him aware of the importance of the area. A comprehensive discussion of the whole field of methods improvement would be far beyond the scope of this book, because it involves principles of tool improvement, machine design, and other matters that are not psychological in nature.

---

[14] F. B. Gilbreth, *Bricklaying System* (New York, M. C. Clark, 1909).

The principles of motion economy have been divided into three types,[16] namely:

I. Principles of motion economy as related to the use of the human body.

---

[15] This field, although basically a part of experimental psychology, has been largely developed in the United States by industrial engineering rather than by psychology, owing in large measure to the fact that when productive industry first began to recognize the importance of motion and time study as a tool of scientific management, American psychologists, with very few exceptions, were interested in other fields and did not develop this area of psychotechnology. The outstanding exception may be found in the work of Frank and Lillian Gilbreth. Mrs. Gilbreth, a professional psychologist, clearly saw the possibilities for applied psychology in this field, and with her husband, a professional construction engineer, published a series of papers and books in which were formulated most of the principles of motion study that are most widely in use today.

In Great Britain the development was somewhat different. The National Institute of Industrial Psychology recognized motion and time study as a legitimate field of psychology and undertook the development of the area that, in America, went by default to industrial engineering.

[16] This classification is taken from R. M. Barnes, *Motion and Time Study* (3rd ed.; New York, John Wiley & Sons, Inc., 1949), Chapter 15.

II. Principles of motion economy as related to the management of the work place.

III. Principles of motion economy as related to the design of tools and equipment.

Motion Economy as Related to the Human Body. The first of the above basic principles is particularly pertinent in the study of human work activity. Under this principle, a number of secondary principles have been developed.[17] The first three are as follows:

1. The two hands should begin as well as complete their motions at the same time.

2. The two hands should not be idle at the same instant except during rest pauses.

3. Motions of the arms should be made in opposite and symmetrical directions and should be made simultaneously.[18]

Numerous experimental demonstrations of the time saved by following these principles have been made. Several such experiments are reported by Barnes,[19] typical of which is the job of assembling a bolt and three washers. On this job, in the study reported, it was found that a 53 per cent increase in production resulted from a work rearrangement to permit adherence to the three principles listed above.[20]

The fourth principle reads:

4. Hand motions should be confined to the lowest classification with which it is possible to perform the work satisfactorily.

The five general classes of hand motions include, first, the fingers by themselves, with the others involving, progressively, other body members in this order: wrist, forearm, upper arm, and shoulder. Generally speaking, lower classifications require less time and physical effort than the higher classifications. There are, however, exceptions to this generalization. Controlled investigations have shown, for example, that for certain types of activity, wrist motion[21] and forearm motions[22] can be made more rapidly and with less fatigue than finger motions. In another investigation it was found that, at least under certain conditions, wrist and elbow movements are faster than finger or shoulder movements.[23]

---

[17] The student who is interested in studying these in more detail and also in studying principles falling under Class II and Class III is referred to *ibid.*, Chapters 15 to 17.

[18] These three principles were first formulated by F. B. and L. M. Gilbreth. "A Fourth Dimension for Measuring Skill for Obtaining the One Best Way to Do Work." *Society of Industrial Engineers Bulletin*, 1923, 5, 6.

[19] Barnes, *op. cit.*, 194–198.

[20] *Ibid.*, 197.

[21] M. Smith, M. Culpin, and E. Farmer, "A Study of Telegraphers' Cramp," *Industrial Fatigue Research Board, Report No. 43* (1927).

The fifth principle:

> 5. Momentum should be employed to assist the workers whenever possible, and it should be reduced to a minimum if it must be overcome by muscular effort.

This principle was used in F. B. Gilbreth's early studies of bricklaying.[24] Methods were developed that made use of the momentum of the moving brick in forming the mortar into the joints.

The sixth principle:

> 6. Smooth continuous motions of the hands are preferable to zigzag motions or straight-line motions involving sudden and sharp changes in directions.

One investigation has shown that 15 to 25 per cent of a work cycle is used in changing hand direction when an abrupt change in direction is made.[25] If the work layout can be arranged to eliminate this kind of lost time, a marked saving in time is effected.

The seventh principle:

> 7. Ballistic movements are faster, easier, and more accurate than restricted (fixation) or "controlled" movements.

Although there is little experimental evidence on this subject, the experience of industrial engineers indicates that ballistic movements, in comparison with controlled movements, are more powerful, more accurate, faster, and less fatiguing.[26]

> 8. Rhythm is essential to the smooth and automatic performance of an operation, and the work should be arranged to permit an easy and natural rhythm wherever possible.

Rhythm has been defined in at least two ways—as a repetitive cycle of motions, and as a sequence of accented motions that may or may not be repetitive.[27] The major characteristics of rhythmic action, in the latter sense, are rapid movements through an arc, with a sudden feeling of termination at the end of the movement. Rhythmic motions, therefore, involve more ballistic than controlled motions.

---

[22] R. H. Stetson and J. A. McDill, "Mechanisms of the Different Types of Movement," *Psychological Monographs*, 1923, **32**, 37–40.

[23] W. L. Bryan, "On the Development of Voluntary Motor Ability," *American Journal of Psychology*, 1892, **5**, 125–204.

[24] F. B. Gilbreth, *Motion Study* (New York, D. Van Nostrand Company, Inc., 1911).

[25] Barnes, *op. cit.*, 220.

[26] *Ibid.*, 224–226.

[27] R. H. Stetson, "A Theory of Rhythm and Discrete Succession," *Psychological Review*, 1905, **12**, 258–270.

Measuring Effects of Methods. Principles of motion economy such as those mentioned above have found wide acceptance throughout industry as a basis for changing (i.e., improving) methods of work. There have been many circumstances in which the methods developed from such principles have been found, through objective study, to be superior in some respect to original methods. Some examples of such studies were mentioned above, in the discussion of certain of the principles listed.

A rather unique study of the effects of a change in method of work is described by Brouha.[28] The work in question involved, in part, heavy lifting and the use of a comparatively heavy hammer above shoulder level. Partial mechanization of the work eliminated these aspects of the job. Heart rate recovery curves were used in comparing the physiological stress before and after the change in method. The "before" and "after" curves are shown in Fig. 16.11. The difference between these indicated very clearly the reduced physiological costs of the new method.

Generally speaking, the evidence from most pertinent investigations tends to support the general principles and procedures of methods improvement. It should be pointed out, however, that the application of a set of methods-improvement principles does not in all cases insure the development of an optimum method.

A study reported by Lauru,[29] for example, raises some question about the general applicability of such principles. This report cites research on work activities with the Lauru platform, which was described in Chapter 15 (p. 438). A comparison was made with the Lauru platform of the physical forces created by the movements of a bricklayer working by three different methods, as follows:

1. The conventional method.
2. The "improved" method developed by Gilbreth,[30] mentioned earlier in this chapter; this method involves particularly the principle

---

[28] L. Brouha, "Fatigue—Measuring and Reducing It," *Advanced Management*, January 1954, **19**, 9–19.

[29] L. Lauru, "The Measurement of Fatigue, Part 2," *Manager*, 1954, **22**, 369–375.

[30] Gilbreth, *op. cit.*

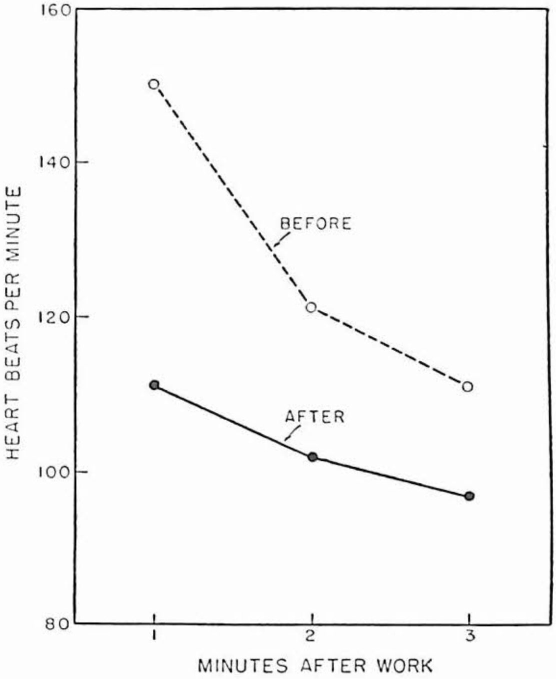


FIG. 16.11. Heart rate recovery curves for manual job before and after partial mechanization. (From Brouha, op. cit.)

of bimanual *symmetry* of movements (moving the arms in opposite and symmetrical directions—principle No. 3).

3. A new method involving the use of only one hand at a time (thus departing from the principle of symmetry).

The forces resulting from these three methods are shown in Fig. 16.12. In studying these, it should be kept in mind that the variability of the curves above and below the base line is indicative of the magnitude of the forces created by the worker during the task. It will be seen that there is noticeably less variation with the new method than with the others, including Gilbreth's, thereby suggesting that the new method requires less physical exertion than the others.

It should also be added that the new method, aside from requiring less physical exertion, also took less time; there was a saving of 18.4 per cent in time over the original method, and of 6.4 per cent over the Gilbreth method.

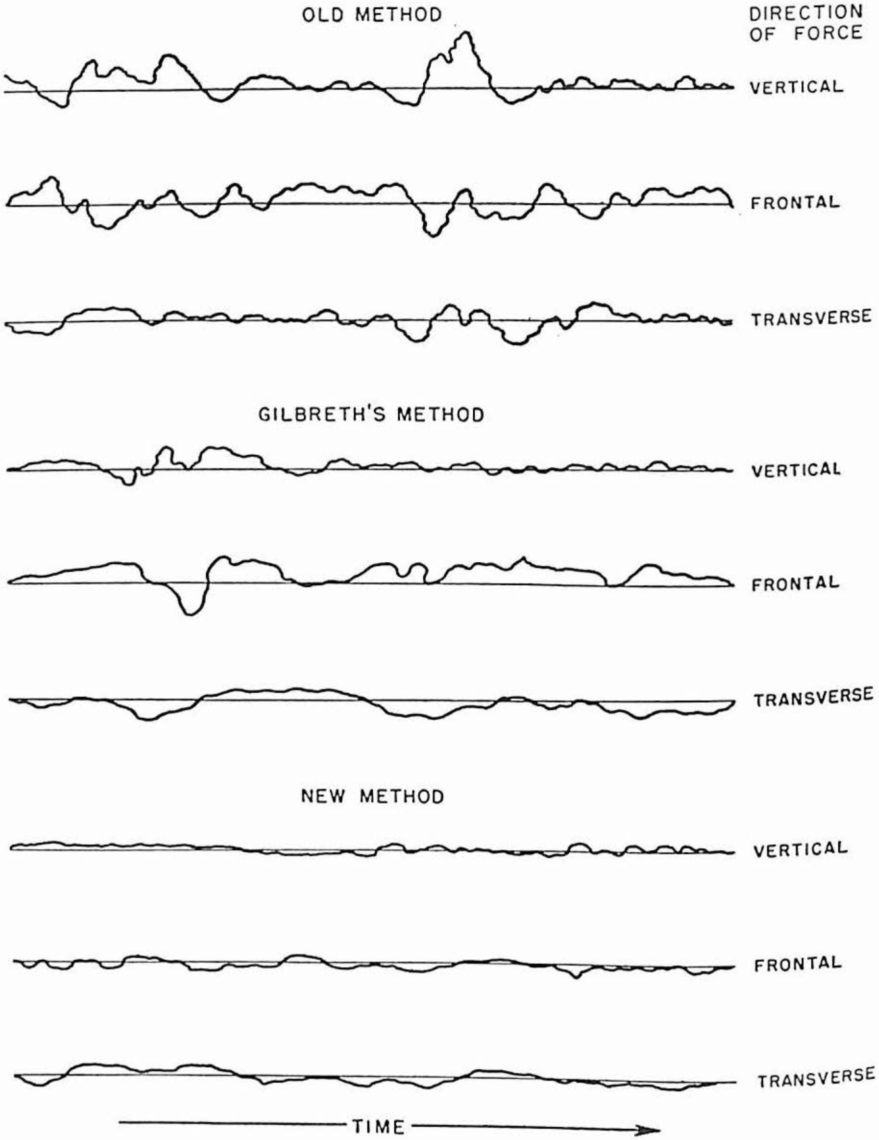


FIG. 16.12. Physical forces exerted when laying brick by three different methods. Forces were measured by use of the Lauru platform in three directions, namely, vertical, frontal, and transverse. The third bricklaying method required the least exertion. (Adapted from Lauru, op. cit.)

In particular, this study seems to cast some doubt about the general validity of the principle of bimanual symmetry. This should not be taken to suggest that this principle has no applicability in methods improvement, but perhaps it does suggest that further research is needed to determine the work circumstances under which the principle is, and is not, applicable. Perhaps other principles of motion economy also could well be subjected to similar research.

17

WAGES AND JOB EVALUATION

It is unnecessary to point out to the readers of this book that wages constitute one of the most persistent problems in the whole field of labor-management relations. Since unions first expressed their demands, an insistence on a higher general wage level as well as on certain other wage adjustments has characterized a majority of labor-management disputes. The importance of monetary demands by organized labor has not lessened in recent years, even though more nonmonetary demands are being made now than in the past. An analysis of 2,055 union demands made in a sample of 821 dispute cases coming before the National War Labor Board during the years 1942–44 revealed that 910, or 44 per cent, of the demands involved monetary issues. A partial summary of this work has been published by Tiffin and Lawshe[1] and is reproduced in Table 17.1. The results tabulated show something of the variety of ways in which monetary demands may be made. The results also show that the percentages of "attained" and "partially attained" demands (even during the operation of a nationwide wage "freeze") are sufficiently high to indicate that labor's arguments were considered rather persuasive by the members of the War Labor Board during World War II.

These facts suggest that we should carefully scrutinize the field of psychology (as well as such fields as economics and sociology) in the hope of finding a body of fact, or perhaps *a technique* or *method*, that will help us to arrive at a fair solution of labor-management monetary disputes. Unfortunately, there does not seem to be any single *method*, acceptable to both labor and management, that may

---

[1] J. Tiffin and C. H. Lawshe, Jr., "War Labor Board Decision Trends," *Personnel*, 1945, 22, 78–83.

TABLE 17.1

National War Labor Board Disposition of 910 Monetary Demands Made by Unions in 821 Dispute Cases Coming Before the Board

Type of Demand	*Per Cent Attained*	*Per Cent Partially Attained*	*Per Cent Attained or Partially Attained*	*Per Cent Refused*
Wages, minimum rates	63	11	74	22
Wages, hiring rates	50	17	67	25
Wages, substandard rates	47	29	76	24
Bonus payments	46	18	64	36
Severance pay	44	13	57	31
Wage adjustments	44	28	72	27
Wages, inequalities	43	26	69	28
Wages, going wage rates	42	22	64	36
Wages, cost-of-living adjustment	38	29	67	30
Wages, incentive wage system	37	23	60	37
Wages, trainee rates	36	28	64	36
Premium wage rates	34	27	61	38
Wage differentials	31	33	64	36
Wages, job rates	23	15	38	54
TOTAL	41	26	67	31

be applied to give a clear-cut solution to all wage disputes. The problem is not like one in arithmetic or mathematics, where, in seeking to find the square root of 63, anyone who knows the *method* will arrive at 7.937+ as the answer, because there *is* no method of solution that is generally acceptable to both labor and management. But there are certain factors that have traditionally affected wage rates, and an understanding of some of the most influential among them will help not only in developing a harmonious labor-management relationship but also in showing to both parties the vital importance of accepting some *systematic* method in the setting of wage rates.

Factors affecting wages can be divided into two broad (and sometimes overlapping) classes: those affecting general (i.e., community and industry) wage levels and those affecting wage rates for different jobs within a plant.

FACTORS AFFECTING GENERAL (COMMUNITY AND INDUSTRY) WAGE LEVELS

Supply and Demand. When no governmental or social restrictions on wages are in effect, the law of supply and demand has traditionally exercised a marked effect on wage levels. When labor is abundant, wages decrease; when labor is scarce, wages rise. Many managements that formerly allowed the operation of this principle to affect their wage structure have recently become more and more reluctant to follow the dictates of this "law." By paying an extremely high wage at times when labor is scarce, management has found that it has had trouble in meeting cost standards, and has been faced with a real difficulty in explaining the reason for wage reductions when the inevitable cut-back becomes necessary. And conversely, it has found that the hidden costs in the form of employee dissatisfaction and pent-up resentment that result from the payment of a low wage when labor is abundant often overbalance in the long run the savings in wages. Basically, of course, a wage program must result in pay levels that have a realistic relationship with those in the general labor market area.

Governmental Wage Controls. Governmental wage controls fall into two major types: minimum rate and maximum rate restriction. The first type is illustrated by the Fair Labor Standards Act of 1938 and amended, providing for higher minimum wages, in 1949 and again in 1956. This law provides that certain minimum wages must be paid by all companies engaged in interstate commerce or in the production of goods for interstate commerce. The setting of maximum wage rates by governmental order is illustrated by the directives of the Wage Stabilization Section of the National War Labor Board during World War II and by the directives of the Wage Stabilization Board during the period of production for defense beginning in 1951. During such periods, maximum wage rates are set to minimize the monetary inflation that inevitably would result if wages are unlimited while much of the productive facilities of the country are engaged in production of materials for war or defense rather than for civilian consumption.

Civic Relations of the Company. Some companies have traditionally paid a higher wage than the prevailing rates in their locality in order to foster the continued good will of workers in the community. A company cannot, of course, follow this policy unless its competitive position in the manufacture of its product will permit such expenditures for wages. Companies that have been able to follow this policy have frequently built for themselves a *status* in the community that has minimized certain problems in employment experienced by other companies in the locality.

Cost of Living. This is a factor that has been considered systematically by some companies in determining general wage levels. Yoder[2] discusses the use of cost-of-living indexes for this purpose, and mentions several companies that employ this method of wage adjustment. He also analyzes a number of cost-of-living indexes and discusses the advantages and limitations of each as a basis for general wage adjustments.

Current union-management contracts often include provision for wage adjustments based on a cost-of-living index.

Regional and Industrial Differentials in Workers' Earnings. Closely related to figures on cost of living are variations in the general level of workers' earnings in different communities or localities. Workers living in cities where rents are relatively high, the cost of transportation to and from work is great, and other incidental expenses are proportionately large, are traditionally paid, by most companies, a higher general wage than workers with the same skills employed in smaller and more rural areas. This tendency has, at times, caused rural employees to feel that they are underpaid. In terms, however, of "real wages," i.e., what the money will buy, it is sometimes true that urban workers are paid less in "real wages" in spite of their greater monetary wages.

Strength of Organized Labor. When labor is organized and is directed by strong leadership, increases of general wage levels are sometimes obtained as a result of union-management negotiations. This statement does not mean that the strike, as such, is necessarily a device that should be thought of as the sole factor in determining wage levels. Any or all of the factors mentioned previously may operate as influencing factors, whether with or without the presence of organized labor. Nor does it mean that only through the strike is a responsible union leadership likely to exert any influence upon the wage policies of a company. The mere process of modern collective bargaining, as described by Hill and Hook,[3] sometimes results in a degree of understanding by management of labor's point of view that brings about certain wage adjustments, just as it sometimes results in a degree of understanding by labor of the problems of management that causes a withdrawal of certain wage demands.

---

[2] Dale Yoder, *Personnel Management and Labor Relations* (Englewood Cliffs, N.J., Prentice-Hall, Inc., 1943), 420–426.

[3] Lee H. Hill and C. R. Hook, Jr., *Management at the Bargaining Table* (New York, McGraw-Hill Book Company, Inc., 1945).

Cost of Production. Whatever may be a company's point of view toward any or all of the factors that may have an effect upon the general wage level, the cost of production in relation to selling price must, of necessity, always set a ceiling on the wages that can be paid by any company that intends to remain in business. No company can follow for long a wage policy that requires a labor expenditure of $1.10 on a commodity that must be sold (either because of competition or because of price ceilings) for $1.00. Obvious as this limiting factor may appear, it has not always been kept properly in mind by certain groups that have attempted to set wages in terms of one or more of the factors previously discussed without considering the selling price of the product.

The factors mentioned above have been discussed to bring out the fact that many elements, several of which are not an immediate concern of the psychologist as such, exercise an important influence in determining general wage levels. In any specific plant situation several of these factors usually operate simultaneously, and are of varying relative importance depending upon the conditions operating at that time.

FACTORS AFFECTING WAGE RATES WITHIN A PLANT

In addition to the factors that influence wages in general, certain additional factors operate predominantly to affect the relative wages of men or women on specific jobs within the plant.

Incentive vs. Hourly Paid Jobs. The earnings of employees working under an incentive wage plan (which provides that they be paid in proportion to their production) are often different from those of employees on the same or equivalent jobs who are paid by the hour. Indeed, incentive plans of wage payment are frequently installed to increase production, an increase that typically results in concomitant increases in the earnings of the employees. The installation of a suitable incentive plan for any given job or group of jobs is a matter of primary concern to the industrial engineer rather than to the industrial psychologist. Many variations of the basic incentive principle are found in the incentive plans of wage payment used by modern industry. Yoder[4] has summarized and described a number of the more widely used incentive systems. A discussion of these plans would go beyond the scope of this book.

---

[4] Yoder, *op. cit.* 390 ff.

Nonfinancial Incentives. Certain jobs are considered by many employees to be desirable for social or similar reasons, rather than for the wage paid to employees on these jobs. When this situation exists it is not uncommon to find the rates for such jobs set somewhat lower than are those for other jobs that require approximately the same level of ability, experience, or skill. Employees are often willing, or even anxious, to be assigned to such jobs because of the recognition or status involved, even though the actual monetary return is somewhat less than that from other jobs. Under the operation of any standard job evaluation system, such as will be described later, the importance of this factor in determining rates is usually minimized or even eliminated, but in the absence of a job evaluation system it sometimes exercises a marked effect upon certain rates within a plant.

Discrimination against Women and Other Groups. Until quite recently certain plants followed a wage policy that permitted different rates to be paid to women on the same, or approximately the same, job as men. This practice was usually justified by the belief that women, because of their limited strength and physique, were unable or should not be asked to do all of the things that men on the same jobs were accustomed to do. With the passage of recent social legislation and with the increased use of standard job evaluation systems, however, it has become necessary that a real difference in job requirements be demonstrated before a different rate can be set for men and women on the same job. Many inequalities formerly justified on this basis have therefore been eliminated by modern industry, and different rates for men and women for the same type of work are now less frequently encountered.

Company Policy. A company going through the period of development intervening between a small organization and a large one usually finds itself compelled to place an ever-increasing number of workers on a wide variety of newly created jobs for which specific and equitable rates could not have been set in advance. In this stage of a company's development, there is seldom a single man or a department charged with the responsibility of developing and maintaining an all-inclusive wage structure that covers various employees on various jobs and in various departments. Many different supervisors in different departments hire men or women for various jobs, each one establishing within his own department rates that he considers fair and reasonable. Almost invariably the result of such a procedure is a set of rates for different jobs that are not properly and fairly related to one another. As the company grows and as these rates become the traditional rates by being paid over a period of time, it is not uncommon to find attempts made to justify the rate schedule on the grounds that the rates are in conformance with "company policy." This is only one of many illogical (and sometimes indefensible) practices that arise during the period of a company's growth and that must be remedied when that company reaches a size that requires integration and co-ordination of activities in various departments. A systematic job evaluation program, of the type that will be described later in this chapter, results in the most reasonable and adequate solution of this problem.

Supply and Demand. In addition to the effect that labor supply and demand have on the general wage level within a plant, these factors, at times, have influenced the rates paid for specific jobs. For example, if a number of welders are needed and there are few men available who are qualified for this work, the welder rate has sometimes been raised in an attempt to get more men for the job.

This method of "solving" a labor shortage in any particular classification is really no solution at all. It does not create more qualified employees. At best, it can do no more than "drain" men from other plants in the locality, and these other plants, by adopting the same procedure, can rapidly "drain" them back. In a very short time such an interplay will raise the rate for the job to an unreasonable level—unreasonable in terms of labor cost and also in comparison with other jobs requiring equivalent experience or skill. This latter point is perhaps the most persuasive reason for not allowing temporary changes in supply-and-demand factors to affect rates for specific jobs. Once a systematic and equitable rate structure has been installed, rates within established brackets should never be changed without a formal re-evaluation of the job. When fewer men for a given classification are available than are needed, it is advisable, both from the point of view of economy as well as from that of sound industrial relations practice, to institute a training or upgrading program to qualify men to fill the vacancies.

Collective Bargaining Agreement. Union contracts sometimes cover specific rates or rate ranges for specific jobs. This provision may result from the activities of a joint labor-management committee on job evaluation, or from a union demand for certain rates for certain jobs when there is no formal job evaluation program in operation or when a program is in operation but the union has not been asked to participate in it.

A company can usually eliminate unreasonable union demands with respect to certain jobs by enlisting active union participation in the job evaluation program.

Job Evaluation. A systematic evaluation of all jobs within a company is rapidly becoming the most widely accepted method of setting satisfactory and equitable rates or rate ranges for various jobs. Job evaluation is a subject to which several methods and techniques of psychology have been extensively applied. This topic will therefore be discussed in some detail in the section to follow.

One of the purposes of job analysis, as described briefly in Chapter 3, is to form the basis for setting equitable rates for various jobs. This evaluation process is called job evaluation. Job evaluation is the rating of jobs according to a specifically planned procedure in order to determine the relative worth of each job.

There are many reasons for the use of a systematic plan in setting wage rates. Without such a plan rates will be affected by many of the previously discussed factors in ways unrelated to one another. Under such circumstances, the rate structure will contain many inconsistencies caused by rates that are not "in line." With a job evaluation plan in operation, inconsistency in rates is minimized and the entire wage structure becomes unified.

The installation of a job evaluation program involves several steps. Some of these steps are required by the mechanics of any system of job evaluation. Other steps, affecting primarily the industrial relations aspects of the installation, have been found expedient by various companies.

GENERAL PROCEDURES IN INSTALLING AND OPERATING A JOB EVALUATION PROGRAM

In installing and operating a systematic job evaluation program, several aspects of the project should be carefully considered in advance. Proper steps should be taken to anticipate difficulties, or at least to minimize them if they arise.

The Job Evaluation Committee. Job evaluation requires not only technical competence in the field (including full knowledge of the system to be used), but also the active and enthusiastic co-operation of men representing several departments of the company. Most closely concerned are the industrial relations, personnel and employment, industrial engineering, and production departments. The job evaluation committee charged with the responsibility of installing the system should therefore be made up of men representing the interests of each of these groups. As to the advisability of having formal representation of organized labor on the job evaluation committee, there are different feelings by different managements. There is, however, rather general agreement that union representation on the job evaluation committee is advantageous so long as the recommendations of the committee are advisory and the committee is not given final authority on wage matters.[5] The basic reason for providing for union representation is that union members will very naturally want to know "what is going on." Unless some provision is made for their receiving direct, accurate, and comprehensive information about the project, they are apt to arrive at an entirely erroneous conception of its purpose. They may conclude, for example, that the jobs are being studied for the purpose of "cutting wages." It is obvious that the program will not function properly if initiated under such circumstances. As Daniels[6] has pointed out, controlled participation as a means of gaining employee acceptance and understanding of a job evaluation plan implies that management takes—and keeps—the initiative in the sense that it sets the objectives of the participation. Participation may take a variety of forms depending upon company size, time available, and other limitations. But regardless of the *form* of the participation, the use of the *method* will be an asset to the workability of the plan and will contribute to a reduction of the problems that necessitated the plan's introduction.

Obtaining the Cooperation of Supervision. In the past, the foreman has had many of the duties and responsibilities that have now been taken over by other representatives of management. Among these functions is the setting of wage rates. As a result, and as in the case of other changes in function and duties, many supervisors do not always recognize the necessity of a unified wage structure because the very nature of their respective jobs makes it difficult for them to see the problems of the company as a whole rather than just the problems of their own departments. It is therefore wise for managements to devote as many supervisory sessions or conferences as necessary to discussions of the job evaluation program, to a full explanation of how the system works, and to the relation of the plantwide program to the needs of various departments.

---

[5] Hill and Hook, *op. cit.*

[6] H. W. Daniels, "Winning Acceptance for the Job Evaluation Plan: Controlled Participation Pays," *Personnel*, 1953, **30**, 30–33.

Job Descriptions for Job Evaluation. The first task of the job evaluation committee is to obtain adequate job descriptions. The importance of the latter as a basis for personnel specifications has been emphasized in Chapter 3. Adequate job descriptions are also of vital importance for a system of job evaluation. Jobs cannot be evaluated properly unless their nature is fully known. Just as job titles by themselves are of little or no value in writing personnel specifications, so also job titles must be accompanied by full job descriptions when the jobs in question are to be evaluated. Moreover, a given set of job descriptions is not always suitable for the various types of job analysis discussed in Chapter 3 unless care has been taken to include all information required for the different uses to be made of the descriptions. Therefore, when the job descriptions are prepared, it is desirable to know (*a*) what they are to be used for, and (*b*) what information about the jobs is needed for all of the proposed uses. When this information is secured *before* the job descriptions are prepared, the time and cost of doing this work over at a later time are usually saved.

Evaluation of Jobs. Upon completion of the job descriptions, the jobs are then evaluated by whatever system has been developed or adopted by the committee. The various types of systems will be described later. This evaluating process (with most systems) results in the assignment of numerical values, such as points, to individual jobs.

Converting Evaluations to Money Values. The next basic process in carrying out a job evaluation program is that of converting the evaluations to money values (hourly wages or salaries). A later section will be devoted to this process, so we will not discuss it in detail here.

Provisions for Evaluation of New Jobs. No matter how thorough and satisfactory a job evaluation program is when it is first installed, provision must be made for continual, careful study and evaluation of new jobs as they are created. Jobs are not static because work processes are not static. The job evaluation committee should be a standing committee, and it should meet at regular intervals in order to keep the evaluations current with the creation of new jobs and elimination of old ones as work processes change.

Getting Jobs in Line after the System Is Installed. A problem that is invariably encountered when a job evaluation system is installed for the first time is that some jobs will be found "out of line." These jobs will be of two types—those that have previously had too high a wage rate and those that previously have been set too low.

No problem is encountered in connection with the latter jobs, because the rate increase indicated by the evaluation will, of course, be accepted by the employees on these jobs. It is just as important, however, to fit those jobs that previously have been overpaid into the basic structure. It is obviously unwise, for many reasons, to adopt the policy of cutting these rates abruptly, and such a policy should never be followed (and almost never is followed) by a plant management. Two procedures are recommended for getting overpaid jobs back "in line." First, every effort should be made to upgrade employees on such jobs to other jobs that permit them to maintain their present rate but that oblige them actually to earn it. Such upgrading may at times require special training. When this is the case, it is advisable for the company to institute the training and make the upgrading as rapidly as possible.

A second step is to hire any new workers needed for these jobs at the rate determined by the job evaluation. If this second step is taken without the first, however, different employees on the same job will be paid at different rates. Such a situation should be eliminated as rapidly as possible, and should not be allowed to exist at all except for a very short time while the transition is being made.

JOB EVALUATION SYSTEMS

There are several kinds of job evaluation systems in successful operation. These systems differ from one another in certain respects and it is always advisable, if at all possible, to decide on the specific system that is to be used *before* the job descriptions are prepared.

By deciding on the system in advance, particular care can be exercised to include in the job descriptions whatever specific information may be required by the job evaluation system to be used.

Job evaluation systems have been classified in several different ways. Lytle[7] points out that they may be classified according to the *method* and also according to the *technique of measurement* used. This dual division of systems is perhaps more elaborate than is needed to summarize some of the basic plans now in use. We will therefore discuss certain plans that fall under five general groupings, as these cover the more widely used plans now in operation.

Scaling of Jobs by Comparing Entire Jobs (Ranking Method). The simplest application of this method is through the principle of ranking. Under this plan, all jobs are ranked and placed in a continuum from highest to lowest. The ranking should be done by a person or committee familiar with all of the jobs. In a large plant, where several hundred jobs must be evaluated, this necessary familiarity is seldom possessed by any single person, or even by a group of persons. This is one of the most serious obstacles encountered in the use of the ranking system. It can be minimized to some extent by having each rater rank only those jobs with which he is thoroughly familiar. Then, by a comparison of rankings given jobs common to several raters' lists, it is usually possible to establish key points in devising a scale that will include all jobs.

The reliability of rankings can be increased by using various psychological techniques, the value of which has been proved. Among these are the method of paired comparisons (in which the rater compares each job with every other job) and the pooling of rankings of several raters to obtain average rankings.

Scaling of Jobs by Comparing Components of Each Job (Factor Comparison Method). This method is ordinarily known as the "factor comparison method." It has been described in detail by Benge, Burk, and Hay.[8] In this method, fifteen or twenty tentative "key jobs" are first selected. These are jobs that have present rates not subject to controversy and that are considered by the job evaluation committee to be neither underpaid nor overpaid. These jobs are compared with respect to factors common to all jobs. The factors used in the Benge, Burk, and Hay system are:

---

[7] C. W. Lytle, *Job Evaluation Methods* (New York, The Ronald Press Company, 1946). A comprehensive discussion of job evaluation methods is also given by J. L. Otis and R. H. Leukart, *Job Evaluation* (2nd ed.; Englewood Cliffs, N.J., Prentice-Hall, Inc., 1954).

[8] E. J. Benge, S. L. H. Burk, and E. N. Hay, *Manual of Job Evaluation* (4th ed.; New York, Harper & Brothers, 1941).

Mental requirements.
Skill requirements.
Physical requirements.
Responsibility.
Working conditions.

The "key jobs" are ranked in order on each of the factors mentioned, and all appear in each of the lists. The rankings are made independently by approximately ten raters, and are made three times by each rater, with periods of approximately one week intervening between each rating.

The going rate (salary or hourly rate) is then divided for each of the tentative key jobs into the amount being "paid" for each of the factors covered by the factor comparison system. This division of the going rate is done without reference to the previous ranking of the tentative key jobs on the several factors.

From these two independent procedures two rank orders of the tentative key jobs on each factor are determined. The first ranking results from the direct ranking of the jobs on the factors. The second comes from the ranking of the monetary values that have been allocated to each job on each factor following the division of the total rate for each job into the amounts presumably paid for each factor. Any of the tentative key jobs that do not come out with essentially the same rank orders in the two independent ranking procedures are eliminated from the list of key jobs. The jobs remaining constitute the framework of the factor comparison system for the company making the installation. All other jobs in the plant are compared with these key jobs, and each is located in its appropriate place on each of the factors included in the system. The amounts to be paid the job for the various factors are then added, which gives the evaluated rate for the job.

Comparing Entire Job with a Predetermined Scale (Classification Method). In the operation of this method, a series of job-level classifications or brackets is set up by management or by the committee charged with this function. Each job is then assigned to one of the grades. The system is illustrated by such classifications as junior clerk, senior clerk, principal clerk, and so forth, used in the Civil Service system. In operation, this plan usually results in the grade location of each job being determined largely by the rate that it carried before the job evaluation was made. The plan thus involves a tendency to perpetuate any such inequalities as may have existed at the outset.

Comparing Job Elements of Each Job with a Predetermined Point Scale (Point Method). This plan, in one of its many variations, is the most widely used system of job evaluation. A survey of 72 companies having from 100 to 43,000 employees conducted by Smyth[9] in 1952 showed that 81 per cent were using a point scale, 13 per cent the factor comparison system, and 6 per cent other kinds of systems. The percentages were approximately the same for manual, clerical, and executive positions.

Point scales are well illustrated by the system used by the National Electrical Manufacturers Association (ordinarily referred to as the NEMA system). The same system is also used by the National Metal Trades Association. This system requires that each job be studied in terms of each of eleven characteristics, and that one of five degrees of each of these characteristics be assigned to the job. The eleven characteristics with the points corresponding to each degree of each are summarized in Table 17.2.

The meaning of each "degree" is clarified by specific "degree" definitions, which are an integral part of this system. For example, in the National Electrical Manufacturers Association system, the "degree" definitions for the various amounts of experience required by the job are as follows:

Degree	*Amount of Experience*	*Points*
1	Up to three months	22
2	Over three months up to one year	44
3	Over one year up to three years	66
4	Over three years up to five years	88
5	Over five years	110

Similar "degree definitions" are included for the various degrees of the remaining ten factors or characteristics. This plan has been specifically constructed for the evaluation of shop jobs. A similar system, encompassing a different list of factors, has been designed for

---

[9] R. C. Smyth, "Job Evaluation Plans," *Factory Management and Maintenance*, 1952, [110 (1)], 118–121.

TABLE 17.2

Job Characteristics and Point Values Corresponding to Various Degrees of Each Used in the National Electrical Manufacturers Association Job Evaluation System

POINTS ASSIGNED TO FACTORS AND KEY TO GRADES

Factors	First degree	Second degree	Third degree	Fourth degree	Fifth degree
Skill					
1. Education	14	28	42	56	70
2. Experience	22	44	66	88	110
3. Initiative and ingenuity	14	28	42	56	70
Effort					
4. Physical demand	10	20	30	40	50
5. Mental or visual demand	5	10	15	20	25
Responsibility					
6. Equipment or process	5	10	15	20	25
7. Material or product	5	10	15	20	25
8. Safety of others	5	10	15	20	25
9. Work of others	5		15		25
Job Conditions					
10. Working conditions	10	20	30	40	50
11. Unavoidable hazards	5	10	15	20	25

the evaluation of salaried positions; it provides for the rating of salaried jobs on the following factors: education, experience, complexity of duties, supervision received, errors, contacts with others, confidential data, mental or visual demand, working conditions, character of supervision, and scope of supervision. The last two are used only when supervisory duties are involved, which means that all jobs that do not include supervisory duties automatically receive zero points on these two factors.

Under the operation of this kind of plan, each job receives a total number of points (the sum of the points received for each characteristic). These total point values are then used in setting up the monetary wage scale.

Job Description Check List. A method of evaluating clerical jobs reported by Miles[10] involves the use of a check list of activities performed by the job incumbent. It was found in developing this check list that the five operations judged to be most important appeared to be the optimum number to consider for evaluation purposes. For the promotion of good employee relations all operations performed are included but the five judged most important are given the most weight.

---

[10] M. C. Miles, "Studies in Job Evaluation. 9. Validity of a Check List for Evaluating Office Jobs," *Journal of Applied Psychology,* 1952, 36, 97–101.

A definite advantage of this system over the standard point systems in evaluating clerical jobs is that it provides for ease in re-evaluating jobs as they change while a given employee is an incumbent. It is well known that a private secretary may do many things on a job that her successor (or predecessor) was unable to do. As a new girl develops the ability to handle these "extra" activities, it seems only reasonable that her salary should keep in step with her added responsibilities. The check list job evaluation system if applied to each job at periodic intervals (say every six months) makes it possible to keep salaries in line with actual job operations.

CONVERTING JOB EVALUATION RESULTS TO A WAGE SCALE

Whatever job evaluation system is used, the result is a grading of all jobs according to the principles and assumptions involved in the plan. Ordinarily this grading is entirely completed before any reference is made to the wage or monetary aspect of the work, although in some plans the original evaluations are made in terms of cents rather than points. The experience of most men in this field has shown that original evaluation in terms of dollars or cents is inadvisable, because persons on the job evaluation committee are less likely to evaluate jobs on an impartial basis if they think in terms of money values than if they think in terms of points, the money value of which has not yet been determined.

After the evaluations have been completed, however, it is necessary to convert the results into money to develop a wage or salary structure. A simple and generally satisfactory method of making this conversion is through the use of key jobs that are common to various plants in the locality. For example, the jobs listed in Table 17.3 might be found in several plants in a given locality as well as in the plant that is revising its wage structure through a comprehensive job evaluation program. Table 17.3 also gives for each of the "common jobs" the total points for the job and the "going rate" for the job in the locality.

The point values of the jobs listed in Table 17.3 are plotted

TABLE 17.3

Typical Jobs that Might Be Used in Converting Job Evaluation Results to a Wage Structure

Jobs Common to Several Plants in the Locality	*Job Evaluation Points*	*Local "Going Rate"*
Tool inspector	321	1.95
Engine lathe operator	277	1.60
Mason	280	1.84
Millwright	301	1.95
Pipefitter	289	1.89
Punch press operator	204	1.42
Tool maker	344	2.00
Parts wrapper	97	.81
Janitor	110	.88
Laborer	136	1.10

against the "going rates" for these jobs. The result is illustrated in Fig. 17.1. A line is then drawn to show the relationship between points and money. This line may be fitted by the method of least squares[11] if great accuracy is desired. For practical use, however, a line drawn in by inspection of the points is usually sufficiently accurate. From this line one may read directly the appropriate rate in cents per hour for all other jobs that have been evaluated. Many of these other jobs will be unique to the plant in question and will not be found in other plants in the locality. The use of a chart such as the one illustrated in Fig. 17.1 provides for setting appropriate rates for all jobs to form a consistent wage structure.

Inspection of Fig. 17.1 suggests that every slight difference in points results in some difference in hourly rate. In practice, most companies (and unions) feel that the inherent lack of perfect accuracy in the judgments that underlie a set of job evaluations makes it desirable to bracket together jobs of approximately the same point value and to consider these jobs as equal in setting up the wage structure. This bracketing results in so-called *labor grades*. The number of labor grades found in specific wage structures varies from around eight or ten to twenty or twenty-five. The tendency of most current union demands in wage contract negotiations is to favor a relatively small number of labor grades.

When the jobs have been bracketed in labor grades, provision is usually made for wage increases within each labor grade, as illustrated

---

[11] C. C. Peters and W. R. Van Voorhis, *Statistical Procedures and Their Mathematical Bases* (New York, McGraw-Hill Book Co., Inc., 1940).

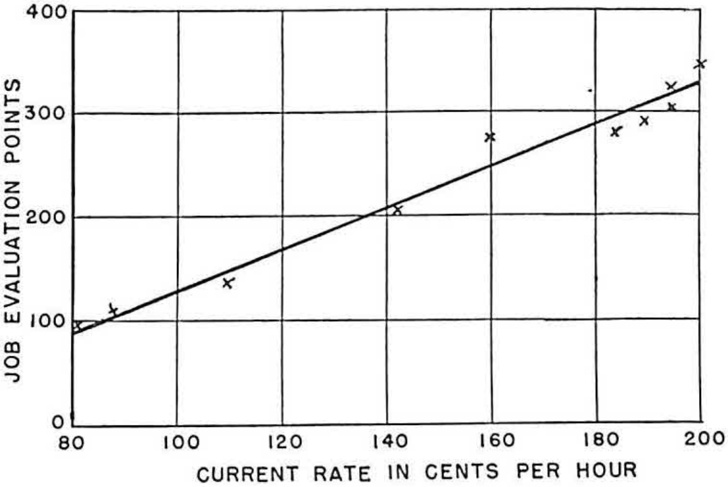


FIG. 17.1. Job evaluation points of typical "key" jobs plotted against the "going rate" for jobs in the locality.

in Fig. 17.2. Various procedures have been used in granting wage increases within labor grades, as well as in upgrading employees to higher categories. Some companies use an automatic acceleration schedule under which specified increases automatically become effective after a specified period of time on the job. This principle is employed most frequently in the lower labor grades and with new

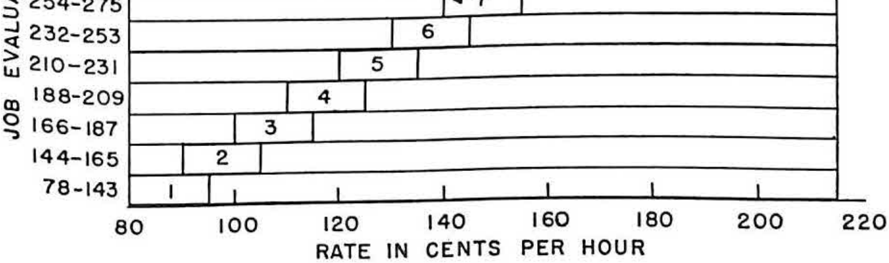


FIG. 17.2. Grouping of jobs into labor grades, with the range of rate of each grade.

value from the maximum values assigned, the actual weightings would not follow the assumed weightings. It is interesting to note, also in Table 17.4, that in three plants the functional weightings computed from variability measures agree rather closely with the weights based on the assigned maximum values. The variable-maximum-value technique of assigning weights, therefore, would seem to be a fairly effective, as well as a simple method of giving any predetermined set of weights to the factors involved.

TABLE 17.4

Weights of Eleven Factors Assumed by Maximum Values of Items in the NEMA System, and Functional Operation of Actual Weightings in Three Plants

Item		Percentage Weights Assumed by Maximum Values of NEMA System	Percentage Weights Operating in System in Three Plants		
			Plant A	Plant B	Plant C
Skill	Education	14	18	17	13
	Experience	22	26	27	26
	Initiative and Ingenuity	14	16	15	14
Effort	Physical Demand	10	6	7	8
	Mental or Visual Demand	5	4	3	5
Responsibility	Equipment or Processes	5	5	5	4
	Material or Product	5	5	3	4
	Safety of Others	5	5	4	6
	Work of Others	5	4	7	4
Job Conditions	Working Conditions	10	8	9	8
	Unavoidable Hazards	5	3	3	8

STATISTICAL STUDIES OF JOB EVALUATION PLANS

Any one of the methods of job evaluation results in a reasonably consistent set of evaluations of the various jobs studied. Each is based upon certain assumptions, which differ from plan to plan, and, as Lytle[14] has pointed out, no one system should be considered to be an employees, but it is also sometimes used at higher levels in the wage structure. A systematic merit-rating program is also used by some plants as a means of identifying employees who are eligible for a wage increase under the prevailing wage structure.

---

[14] Lytle, *op. cit.*, 4.

The particular method that is followed in making wage increases has often been the subject of bitter controversy in union-management contract negotiations. Unions frequently demand rigid acceptance of seniority as the basic determining factor, and the company as frequently contends that proper managerial responsibility requires the granting of wage increases upon the basis of proved merit rather than their automatic occurrence with tenure on the job. Some companies have used job qualification tests for the upgrading of employees, not only with a given labor grade but from one labor grade to another. The Northwest Airlines has in operation such an upgrading program based on paper-and-pencil, as well as performance, tests under terms of a contract with the union concerned, the Air Line Mechanics Association.[12] The use of tests as one of the tools for upgrading offers the dual advantage of assuring members of the bargaining unit that qualified men will be given proper consideration, regardless of their seniority, and of assuring the company that unqualified men will not automatically be given promotions with tenure on the job, irrespective of proved ability.

WEIGHTING OF ITEMS IN JOB EVALUATION

Under a point method of job evaluation such as the NEMA plan, it is usually considered advisable to assign different weightings to the several characteristics. The weights used usually are based upon the experience and judgment of men charged with the responsibility of creating and installing a usable system. In the operation of many plans, different weights are given to the various characteristics by allowing a different maximum number of points for the several items.

In the NEMA system the weights assigned are as summarized in Table 17.4. In actual operation, the functional weightings are determined by the variability of spread of the ratings on each trait.[13] Thus, if the variabilities of the item values were to differ in proportionate

---

[12] *Trade Test Standard Manual for Written and Practical Tests*, 1st ed., 1943, Personnel Department, Training Division, Trade Testing Section, Northwest Airlines, Inc., Minneapolis, Minnesota.

[13] See Appendix A, p. 531.

"exact science." As long as *any* systematic plan is followed that provides for all jobs to be studied in terms of the *same* assumptions and the *same* other stipulations (such as weights for the factors involved), the result is far more satisfactory than a wage structure that has grown piecemeal, without integration. One study[15] has shown that different job evaluation systems give essentially the same results. Chesler found intercorrelations among six different company systems to range from .89 to .93. Two of the systems were factor comparison, two were fifteen-factor point systems, one a thirteen-factor point system, and one a combination ranking and classification system. Using from one to three raters in each plant—raters who were familiar with their respective systems—thirty-five standard jobs were rated with the six different systems. These results show that essentially similar job evaluation results are obtained with different job evaluation systems.

The choice of which system to use is not, therefore, a problem of choosing from among several plans, only one of which can be correct, all the others being wrong. There are practical advantages as well as disadvantages to every system, and any one plan that is installed with care is almost certain to be better than no plan at all. The time and expense involved in installing and operating a job evaluation system, however, are such that it is advantageous to use a method that will accomplish the desired result with a minimum of labor, if the results obtained from a simple system are equivalent to those that would result from a more complex one.

Certain basic methods of psychology are ideally adapted to a quantitative study of job evaluation systems. These methods have been applied in a series of studies reported by Lawshe and his collaborators.[16] Since the NEMA point system (or a system basically the

---

[15] D. J. Chesler, "Reliability and Comparability of Different Job Evaluation Systems," *Journal of Applied Psychology*, 1948, **32**, 465–475.

[16] C. H. Lawshe, Jr., and G. A. Satter, "Studies in Job Evaluation: 1. Factor Analyses of Point Ratings for Hourly-Paid Jobs in Three Industrial Plants," *Journal of Applied Psychology*, 1944, **28**, 189–198.

C. H. Lawshe, Jr., "Studies in Job Evaluation: 2. The Adequacy of Abbreviated Point Ratings for Hourly-Paid Jobs in Three Industrial Plants," *Journal of Applied Psychology*, 1945, **29**, 177–184.

C. H. Lawshe, Jr., and A. A. Maleski, "Studies in Job Evaluation: 3. An Analysis of Point Ratings for Salary-Paid Jobs in an Industrial Plant," *Journal of Applied Psychology*, 1946, **30**, 117–128.

C. H. Lawshe, Jr., and Salvatore L. Alessi, "Studies in Job Evaluation: 4.

TABLE 17.5

Correlations Obtained between Ratings on Selected Factors and Total Point Ratings in Three Plants

Selected Rating Scale Factors	Correlation with Total Points Based on 11 Factors
Plant A:	
Experience (or learning time)	.96
Experience (or learning time) plus hazards	.97
Experience (or learning time) plus hazards plus education	.98
Plant B:	
Experience (or learning time)	.93
Experience (or learning time) plus initiative	.95
Experience (or learning time) plus initiative plus responsibility for the safety of others	.96
Plant C:	
Experience (or learning time)	.86
Experience (or learning time) plus hazards	.91
Experience (or learning time) plus hazards plus initiative	.93

same) has been more frequently used than any other,[17] Lawshe first made certain statistical analyses of the results of several NEMA-point-system installations to determine whether approximately the same results could have been obtained by the use of a simplified system involving fewer characteristics to be rated.

Using a standard statistical technique,[18] a determination was made of which single factor, which two factors, which three factors, and so on, give point ratings that correlated highest with the total point ratings. The results obtained in three different plants using the NEMA system (or a slight modification of it) are summarized in Table 17.5.

In each case the correlations obtained between the most influential

---

Analysis of Another Point Rating Scale for Hourly-Paid Jobs and the Adequacy of an Abbreviated Scale," *Journal of Applied Psychology*, 1946, **30**, 310–319.

C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 5. An Analysis of the Factor Comparison System as It Functions in a Paper Mill," *Journal of Applied Psychology*, 1946, **30**, 426–434.

C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 6. The Reliability of Two Point Rating Systems," *Journal of Applied Psychology*, 1947, **31**, 355–365.

---

[17] Smyth, *op. cit.*

[18] The Wherry-Doolittle Technique, one of the basic statistical methods used for this purpose, is described by Stead, Shartle, *et al.*, *Occupational Counseling Techniques* (New York, American Book Company, 1940), 245 ff.

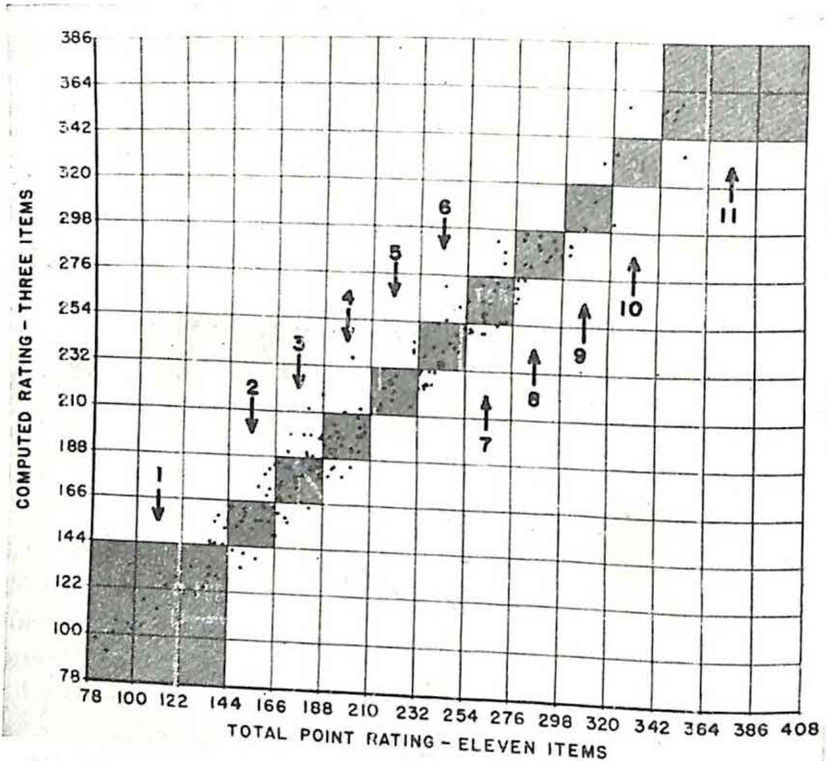


FIG. 17.3. Graph showing ratings computed from three scale items plotted against total point ratings (all eleven items) for 247 jobs in Plant A. The eleven shaded areas define the labor grades designated by the numbered arrows.

three characteristics and the total points determined from all eleven factors were so high that the contribution of the remaining eight factors was negligible.

The facts summarized in Table 17.5 show that *experience* (or learning time) is the most important single factor in determining the total points for the jobs evaluated in all three of the plants. It will also be noted that *hazards* appear among the most important three factors in two of the plants, as does *initiative*. The results further show that three job characteristics give results that, for all practical purposes, are statistically equivalent to those obtained from the use of a more extensive system.

Fig. 17.3 is a graphic representation of the results obtained in Plant A showing the accuracy with which the total points for each

TABLE 17.6

Discrepancies between Total Point Ratings (Eleven Factors) and Ratings Computed from Three Factors for Plant A

POINTS OF DEVIATION	NO. OF JOBS BY LABOR GRADE DISPLACEMENT			
	Same Labor Grade	Displaced One Labor Grade	Displaced Two Labor Grades	All Jobs
0–4	68	9		77
5–9	48	27		75
10–14	20	28		48
15–19	11	21		32
20–24	6	2		8
25–29		4		4
30–34		1	1	2
35–39			1	1
Totals	153	92	2	247

job can be predicted from a knowledge of the points assigned on only *three* characteristics—namely, experience, education, and hazards.

The shaded areas in Fig. 17.3 delineate labor grades. From this chart it is apparent that the vast majority of jobs fall in the same labor grade when evaluated on the basis of the stipulated three characteristics as when evaluated on the basis of all eleven factors.

A summary of the effectiveness of the abbreviated three-factor system in Plant A in terms of number of jobs in which the discrepancy in placement between it and the total point rating is one, two, or three labor grades is shown in Table 17.6.

Table 17.6 shows that only seven jobs deviate from their original placement by more than 24 points, the approximate range of points covered by one labor grade in this system.

In an investigation of the type summarized above, it is possible that a small number of factors can be extracted from the original set so that for the plant installation under investigation the job evaluations yielded by the complete set of factors will be practically duplicated by the abbreviated system, even though the shortened scale would not work equally well as an abbreviation of the same original system in other plants. Otis and Leukart[19] have therefore formulated a question about the general applicability of a shortened scale. They ask: "How would one go about building an abbreviated scale? There is no way of knowing which one of the shortened scales obtained by Lawshe would best apply to a given plant." Otis and Leukart point out that to decide upon one particular shortened scale that will accomplish the same results as a longer scale in any given plant, the longer scale would first have to be used with all of the jobs and statistical analysis then made to identify the specific factors that will work best as an abbreviated scale for that particular plant installation. They state that the savings in time therefore would not be very great.

---

[19] Otis and Leukart, *op. cit.*, 126.

To test whether this criticism is valid, Davis and Tiffin[20] selected three factors, each of which appeared in at least two of the three prediction equations developed by Lawshe,[21] and applied this prediction equation to eight entirely different job evaluation installations. Three of the eight installations were NEMA plans, and five were modified point systems. The three factors used were experience, initiative and ingenuity, and hazards. The correlations obtained between total points and points predicted from the three factors are summarized in Table 17.7. For each of the three NEMA installations the three factors used yielded correlations of .94 or higher with the total points from the entire set of factors. For four of the five point systems that were not NEMA plans the corresponding correlations were .91 or above. For only one of eight installations studied, in Company 5, was the correlation too low to justify predicting total points from the abbreviated scale. This one company was a public utility and was using a job evaluation system that included factors quite different from those included in the NEMA plan.

It would thus seem that an abbreviated system does give essentially the same results as a longer system, and that the short system does not need to be developed specially for every plant where it might be applied.

Salary jobs are ordinarily evaluated by a system that is different from the system used with hourly-rate jobs. The NEMA has a system that includes eleven factors for the evaluation of salary jobs. Lawshe

---

[20] M. K. Davis and J. Tiffin, "Cross Validation of an Abbreviated Point Job Evaluation System," *Journal of Applied Psychology*, 1950, **34**, 225–228.

[21] C. H. Lawshe, Jr., "Studies in Job Evaluation: 2. The Adequacy of Abbreviated Point Ratings for Hourly-Paid Jobs in Three Industrial Plants," *Journal of Applied Psychology*, 1945, **29**, 177–184.

TABLE 17.7

Correlations between Total Points and Total Points Predicted from Three Factors in Eight Job Evaluation Installations

Company	*No. of Jobs*	*Correlation*
No. 1 NEMA	341	.96
No. 2 NEMA	126	.95
No. 3 NEMA	605	.94
Non-NEMA Point Systems		
No. 4, 12 factors	61	.98
No. 5, 11 factors	77	.69
No. 6, 10 factors	185	.96
No. 7, 11 factors	273	.96
No. 8, 23 factors	253	.91

and Maleski[22] have reported a study showing how the evaluations of this eleven-factor system can be practically duplicated by an abbreviated system containing only one, two, or three of the original factors. The correlations between results obtained from a reduced number of characteristics and total point values based upon all eleven factors are summarized in Table 17.8.

In the results for salaried jobs, as in those previously discussed for shop jobs, the experience factor is shown as the most important single characteristic. Also, as in the former instance, ratings based on only three factors are revealed as equivalent to those based on the larger number of factors.

TABLE 17.8

Correlations Obtained between Ratings on Selected Factors and Total Point Ratings for Approximately 400 Salary Jobs

Selected Scale Factors	*Correlation with Total Points Based on 11 Factors*
Experience	.93
Experience plus complexity of duties	.96
Experience plus complexity of duties plus character of supervision	.98

The preceding discussion has been concerned with work on abbreviating the NEMA point systems of job evaluation for shop jobs and salaried jobs. The question may arise whether the same pattern of results would be obtained if another type of point system were similarly analyzed. In answer to this question, Lawshe and Alessi[23] report a study in which the same type of analysis was made of ratings obtained with a point system differing from the NEMA system in that each job is rated on seven characteristics by means of more and finer categories or degrees for each of the factors, and that the point ratings are translated into "rating factors" by a logarithmic conversion chart for the purpose of assigning monetary equivalents. Table 17.9 summarizes the results obtained from this analysis.

---

[22] Lawshe and Maleski, *op. cit.*

From the correlations summarized in Table 17.9 it is apparent that the stipulated three factors give results practically identical with those obtained from the more elaborate system.

After observing the results of statistical simplification of point systems of job evaluation, the question may be raised as to whether corresponding steps toward simplification of other kinds of job evaluation plans will be equally effective. A similar type of statistical analysis of the factor comparison system as installed in a paper mill has been reported by Lawshe and Wilson.[24] Here also the results

TABLE 17.9

Correlation Coefficients between Ratings on Selected Scale Factors and Total Point Ratings Obtained with a Point System Differing from the NEMA System

Selected Factors	*Correlation with Total Points Based on 7 Factors*
Responsibility	.92
Responsibility plus manual skill	.97
Responsibility plus manual skill plus working conditions	.98

TABLE 17.10

Correlation Coefficients between Ratings on Selected Factors and Total Points Obtained with the Factor Comparison System of Job Evaluation

Selected Factors	*Correlation with Original Ratings*
Skill requirements alone	.94
Skill plus working conditions	.97
Skill plus working conditions plus mental requirements	.99

---

[23] Lawshe and Alessi, *op. cit.*

[24] C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 5. An Analysis of the Factor Comparison System as It Functions in a Paper Mill," *Journal of Applied Psychology*, 1946, **30**, 426–434.

show that a reduced number of job characteristics give results that correlate almost perfectly with the total factor evaluations. These correlations are summarized in Table 17.10.

The fact that fewer than the original number of factors are necessary to give essentially identical results is also shown in a study by Howard and Schutz.[25] These men factor analyzed a 12-factor salary job evaluation plan and found that a single fundamental factor corresponding to Lawshe's "Skill Demands" accounted for 99 per cent of the variance in total points. Oliver and Winn[26] report similar results. They found that six factors of an 18-factor point system plan account for most of the variance of the system. The short six-factor system classified 62 of 86 salaried jobs in same salary grade as the 18-factor plan and there was a discrepancy of only one salary grade in the classifications of the remaining 24 jobs.

From the results of the several comprehensive studies reported by Lawshe and others, it seems rather clear that job evaluation results statistically equivalent to those derived from more elaborate systems can be obtained from a very much abbreviated plan.

RELIABILITY OF JOB EVALUATION

Any form of job evaluation is basically an attempt to measure the relative worth of various jobs by a process of individual, pooled, or committee judgments. The ratings obtained (regardless of the system in use) should have a satisfactory reliability, just as the results of any other type of measurement must have a satisfactory reliability, before usable results can be expected. Other things being equal, if two systems have different reliabilities, the one with the higher reliability should always be chosen.

In order to estimate the reliability of the most commonly used point system (the NEMA plan), and also to estimate the reliability of a simplified point system based upon the work summarized in the preceding sections, a series of reliability studies has been conducted and reported by Lawshe and Wilson.[27] In this work, forty job descriptions of commonly known jobs were chosen for evaluation. Twenty men experienced in job evaluation made the ratings. Each job was rated by five men using the NEMA plan, and also by five other raters using a simplified plan that included only four factors. Each rater made his ratings independently. The four factors included in the simplified plan were chosen after factor analyses had been made of several more elaborate systems. These factors were selected to cover as fully as possible the basic characteristics revealed by the several factor analyses.

---

[25] A. H. Howard and H. G. Schutz, "A Factor Analysis of a Salary Job Evaluation Plan," *Journal of Applied Psychology*, 1952, **36**, 243–247.

[26] J. A. Oliver and A. Winn, "An Abbreviated Job Evaluation Plan for Salaried Personnel," *Personnel*, 1951, **28**, 225–229.

[27] C. H. Lawshe, Jr., and R. F. Wilson, "Studies in Job Evaluation: 6. The Reliability of Two Point Rating Systems," *Journal of Applied Psychology*, 1947, **31**, 355–365.

The results of the reliability investigations of the two systems are summarized in Table 17.11. In this table, the factors used in the NEMA system are listed in column (1), in the order of their obtained reliabilities. The reliabilities, listed in column (2), were obtained by computing the coefficient of correlation between the ratings of each rater and those of every other rater for each factor in the list. The average of the correlations thus obtained was taken as the best estimate of the reliability of a single rater on the factor. In column (3) of Table 17.11 are given the "stepped up" reliabilities, i.e., the values

TABLE 17.11

Reliability Data on the NEMA Job Evaluation System and a Simplified Point System Involving Four Factors

NEMA System			Simplified System		
(1) *Factor*	(2) r	(3) r_5	(4) *Factor*	(5) r	(6) r_5
Experience	.82	.96			
Initiative and ingenuity	.78	.95	Learning period	.86	.97
Education	.77	.94			
Responsibility for safety of others	.54	.85	General schooling	.79	.95
Working conditions	.54	.85			
Responsibility for work of others	.51	.84	Working conditions	.61	.89
Physical demand	.47	.82			
Responsibility for equipment or processes	.41	.78			
Responsibility for materials	.40	.77			
Mental or visual demand	.37	.75			
Unavoidable hazards	.34	.72	Job hazards	.51	.84
Total Points	.77	.94	Total Points	.89	.98

that would be expected if averages of the ratings of five raters working independently were used consistently.

The four factors included in the simplified system are listed in column (4). Each factor is located to correspond with the approximately equivalent factor of the longer scale.

Columns (5) and (6) give, respectively, the obtained and "stepped up" reliabilities of the factors of the shorter scale. At the bottom of the table are given the obtained and "stepped up" reliabilities of the total scales.

It will be noted in Table 17.11 that the factors of the shorter scale are all higher in obtained reliability than the corresponding factors of the NEMA scale, and this is true whether one considers the obtained reliabilities or the "stepped up" reliabilities. It will be further noted that the reliability of the total shorter scale is higher than the corresponding reliability of the longer scale.

The greater reliability of the shorter scale may be due in part to the use of more carefully worded descriptions of the degrees of the various characteristics, rather than to the mere fact that a shortened scale is being used. Familiarity with the scale and the experience with its use will not explain the results, because the majority of raters using the NEMA scale had had practical experience with it, whereas *none* of the raters had ever seen the shortened scale before the experiment.

Table 17.11 also shows that the obtained reliability of the total NEMA point ratings is lower than that of several of the individual factors in this system. In the case of the simplified system, the total points resulted in a higher reliability than that obtained for any specific factor.

The study by Chesler,[28] referred to previously, also deals with reliability of job evaluation. In this study seven raters from seven different companies each evaluated the same thirty-five standard jobs from job descriptions. A standard job evaluation manual with twelve factors was used by all raters. The evaluations made by each rater were correlated with those from every other rater. The correlation coefficients ranged from .93 to .99, with an average of .97. These results indicate that different raters, using the same manual and evaluating the same jobs, agree very closely with each other in the job evaluations obtained. These findings also indicate that: (1) The ratings from a shortened system correlate very highly with the ratings of longer systems, (2) The findings are essentially the same for shop jobs and salaried jobs, (3) The effect of shortening the system is the same for point systems and for the factor comparison system, and (4) Ratings obtained from any standardized point system have a satisfactory degree of reliability.

---

[28] Chesler, *op. cit.*

In the light of these findings—that job evaluation systems consisting of only a few factors give essentially the same results as systems involving many more factors, and that a shortened system is at least as reliable as a longer system—it would seem advisable in the future to make greater use of systems considerably shorter than many that are in current use.

SUMMARY

Perhaps the primary implication of this discussion of job evaluation is in terms of emphasizing the utility to management of a systematic, quantitative analysis of information relating to job evaluation systems that are in actual use, or that are being considered for adoption. Such an analysis can provide management with information that can be useful in making more adequate decisions regarding job evaluation programs.

The measurement, quantification, and analysis of data is, of course, central to the principal point of view that the writers have attempted to emphasize throughout this text. In the field of personnel management, as well as in other areas, there are no substitutes for solid, hard facts when making decisions. The systematic application of some of the methods and procedures discussed in this book offers considerable promise of providing pertinent facts about some of the human problems in business and industry that can be used in making decisions directed toward their solution.

APPENDIX A

ELEMENTARY DESCRIPTIVE STATISTICS

DESCRIPTIVE *VS.* SAMPLING STATISTICS[1]

Two basic purposes are served by statistical methods. One of these is to describe a body of data. This is done by means of *descriptive statistics*, which deals with the reduction of a body of data by means of graphic methods, computational methods yielding numerical measures, and tabular methods. The sole aim of *descriptive statistics* is to reduce the original data to charts, graphs, averages, and the like so that the salient facts concerning the data will be more apparent.

The second purpose of statistical methods is to enable the experimenter to learn how safely he can generalize from *descriptive statistics* obtained on a sample. This approach is known as *sampling statistics*. It involves applications of the mathematics of probability and is a very important part of the field of industrial psychology.

This appendix will deal only with certain concepts of *descriptive* statistics, but the importance of *sampling statistics* should not be underestimated and every serious student of industrial psychology should become familiar with these methods.

When many measurements, such as scores or other data, are to be summarized or interpreted, the use of some form of statistical procedure is usually desirable. If a considerable amount of raw data is involved, a simple listing of the data is of little value. Such a listing will not tell us, for example, how the data are distributed, how much they vary, or where in the total distribution they tend to cluster. Further, such a listing is of little value in indicating how the data compare with, or are related to, other sets of data collected under other circumstances. Before a meaningful interpretation of the data can be made it is necessary to reduce them to a *chart* or to *one or two single numbers that may represent the data as a whole.*

---

[1] The distinction between *descriptive* and *sampling statistics* is carefully drawn by John G. Peatman, *Descriptive and Sampling Statistics* (New York, Harper & Brothers, 1947). This book contains a clear and comprehensive treatment of both areas.

GRAPHIC REPRESENTATION OF DATA

The Frequency Distribution and Polygon. A frequency polygon, constructed from a frequency distribution, is a graphic representation of a set of data. The construction and interpretation of a frequency polygon may best be explained by an example. Suppose sixty employees on an inspection job have detected the following number of flaws of a certain type during one week of work:

TABLE 18.1

Number of Flaws Detected by Each of 60 Inspectors During One Week of Work

15	36	40	37	32	13	35	20	33	36	33	16	38	19	33	34	24
36	25	29	27	39	42	31	21	26	28	53	23	51	21	26	39	28
30	31	32	30	29	49	39	30	44	34	37	35	38	35	41	37	43
42	38	45	22	46	41	47	48	34								

From a gross, or even a detailed, inspection of these sixty values one cannot answer such questions as: What is the typical number of defects spotted by an average inspector in a week? How much difference is there between the best and poorest inspectors in spotting defective materials? Is there any preponderance of good, poor, or average ability represented in the performance of these inspectors?

Such questions as these may be answered at a glance if the data are grouped and presented in a chart. One variety of such a chart is a frequency polygon. The steps involved in constructing a frequency polygon are as follows:

1. Determine the range of the values in the raw data. Quickly glance through the data to determine the *highest* and the *lowest* values. The range is the difference between these values. In the case of the sixty inspector records, the highest figure is 53 and the lowest is 13. The range is therefore 53 — 13 = 40.

2. If we find that the range of the data is large (that they are widely spread), it will be more convenient to group them by intervals (class intervals, abbreviated c.i.) with a range in each c.i. of more than one unit. The c.i. is a group of adjacent scores of such a size that from twelve to eighteen c.i.'s cover the range of the whole distribution or all of the data. With a range of forty, a c.i. of 2 would require twenty groups; a c.i. of 3, fourteen groups; and a c.i. of 4, ten

groups. In our illustrative problem, a c.i. of 3 is therefore the proper size to use.

A simple rule of thumb that is helpful in deciding upon the correct size of the c.i. is to divide the range by 15 (15 because, on the average, this is the most desirable number of c.i.'s) and take as the c.i. the whole number nearest to the quotient. In our problem, the range divided by 15 would be 40 ÷ 15 = 2.66. As 3 is the whole number nearest to 2.66, 3 would be the size of the c.i. to be used.

3. Arrange the adjacent c.i.'s in a column, leaving a blank space immediately to the right of this column. The arrangements of the c.i.'s preparatory to the construction of a frequency distribution appear as follows:

TABLE 18.2

Class Intervals to Be Used for Illustrative Data in Table 18.1

51–53
48–50
45–47
42–44
39–41
36–38
33–35
30–32
27–29
24–26
21–23
18–20
15–17
12–14

4. Place a tally mark for each value in the original list of raw data opposite the appropriate class interval. As the first value among the sixty listed in Table 18.1 is 15, the first tally mark should be in the 15–17 c.i. The second value, 36, is represented by a tally mark in the 36–38 c.i. Usually it is advisable to tally the fifth entry in each c.i. with a line across the preceding four tally marks. This simplifies the counting of tally marks at a later time. When all entries have been made, that is, all data tabulated, the frequency distribution appears as in Table 18.3.

5. Lay off appropriate units on squared (cross section or graph) paper so that a graph may be constructed on which the midpoints of the c.i.'s are plotted on the base line and the frequencies or number of cases in each c.i. on the vertical axis. When this is done, the frequency polygon shown in Fig. 18.1 is obtained.

To one familiar with the concept of a frequency polygon, the

TABLE 18.3

Class Intervals, Tally Marks, and Frequencies (f) for Illustrative Data in Table 18.1

Class Intervals (c. i.)	*Tally Marks*	*Frequency* (f)
51–53	//	2
48–50	//	2
45–47	///	3
42–44	////	4
39–41	~~////~~ /	6
36–38	~~////~~ ////	9
33–35	~~////~~ ////	9
30–32	~~////~~ //	7
27–29	~~////~~	5
24–26	////	4
21–23	////	4
18–20	//	2
15–17	//	2
12–14	/	1
		Total = 60

graphic illustration in Fig. 18.1 is a much more meaningful representation of the data than the list of values shown in Table 18.1, or the frequency distribution shown in Table 18.3. The frequency polygon makes apparent at a glance that the typical or average inspector detected around thirty-five defects during a week of work, that the operators range or vary from some who detected only thirteen defects to others who detected fifty-three, and that a majority of the operators are fairly near the average in ability (that is, that not so many are very high or very low as are near the average). In summarizing psychological data it is a definite advantage to be able to present all of the major facts in a single graphic presentation of this type.

In a frequency polygon such as the one shown in Fig. 18.1 the area between the curve and the base line is determined by the number of cases (called N) the graph represents. Thus, a curve portraying 120 cases would cover twice the area of the curve shown (if the c.i.'s are the same for both distributions), and a curve portraying 600 cases would cover ten times this area. This is no disadvantage in many cases, but situations sometimes arise in which it is desirable to keep the total area under the curve the same, regardless of N or the number of cases. To meet this situation we may plot the *percentage* of cases rather than the *number* of cases falling in each c.i. This may be accomplished by dividing each f value in Table 18.3 by the total number of cases in the distribution. Each quotient thus obtained

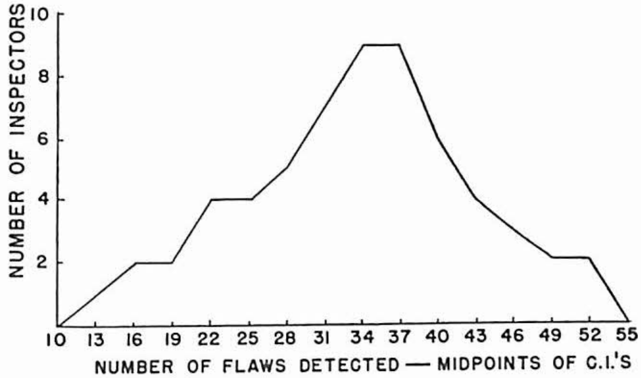


FIG. 18.1. Frequency polygon of illustrative data in Table 18.3.

indicates the percentage of cases from the total falling in the respective c.i. These computations are indicated in Table 18.4.

If the percentages shown in the last column of Table 18.4 are now plotted as the ordinates (vertical axis points) of a frequency polygon, the chart shown in Fig. 18.2 is obtained.

When a frequency polygon is to be compared with a number of other polygons, and when the important facts to be compared deal with the central tendencies and general form of the distributions rather than with the different number of cases plotted in each, the "percentage method" of plotting a frequency distribution is preferred to the "total-number-of-cases method."

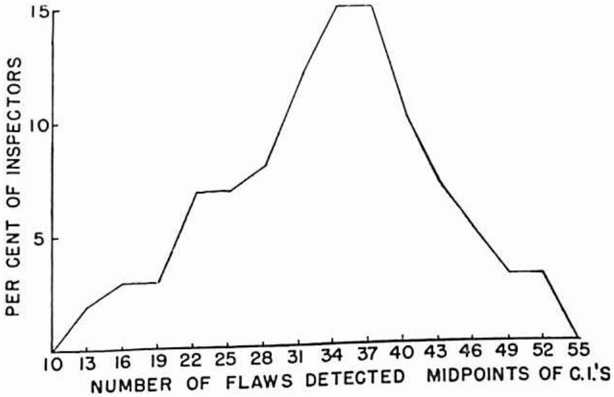


FIG. 18.2. Frequency polygon with ordinates as percentages for illustrative data in Table 18.4.

We may define a frequency polygon, in the light of the above description, as a *curve that portrays data graphically and that is so drawn that the base line represents the varying values of the original data and the ordinates represent the number of cases (or percentage of cases) at each of the raw data values.*

TABLE 18.4

Frequency and Percentage of Cases in Each C.I. for Illustrative Data in Table 18.1

Class Intervals	*f*	*Calculation of Percentage*	*Percentage*
51–53	2	$2/60 = .033$	3
48–50	2	$2/60 = .033$	3
45–47	3	$3/60 = .050$	5
42–44	4	$4/60 = .066$	7
39–41	6	$6/60 = .100$	10
36–38	9	$9/60 = .150$	15
33–35	9	$9/60 = .150$	15
30–32	7	$7/60 = .117$	12
27–29	5	$5/60 = .083$	8
24–26	4	$4/60 = .066$	7
21–23	4	$4/60 = .066$	7
18–20	2	$2/60 = .033$	3
15–17	2	$2/60 = .033$	3
12–14	1	$1/60 = .017$	2
Total =	60		100

The Normal Distribution. The shape of the frequency polygons shown in Figs. 18.1 and 18.2 is typical of the kind of distribution usually found when data obtained from a group of people are plotted. It will be noted that the curves are approximately "bell-shaped," that is, they are high in the center and taper off toward the base line at both ends. If we were to divide the area under such a curve by drawing a perpendicular line from the central high point to the base line, the two parts would be approximately equal in area and would be bilaterally symmetrical in shape. It is well recognized that all, or nearly all, measurements of human traits and abilities result in distributions of approximately this form. Such distributions are called *normal distributions.* A strictly normal distribution conforms to a symmetrical bell-shaped curve that is defined by a mathematical equation, the basis of which is beyond the scope of the present discussion.[2] It will suffice for the beginning student to know that:

1. A normal distribution is bell-shaped, that is, it is high in the center and low at both ends. Its two halves are symmetrical.

---

[2] *Ibid.*, 331 ff.

2. Measurements obtained from a group of persons usually approximate this type of distribution.

MEASURES OF CENTRAL TENDENCY

While the frequency polygon is helpful in giving an immediate graphic description of a set of data, that is, indicating general trends, it is often desirable to present certain quantitative figures that supplement the graphic picture. One of the most useful of such numerical values is a measure of the central tendency of the data. A measure of central tendency may be defined as a single figure or value that is representative of the entire set of data. Three such measures that are in common use are the arithmetic mean, the median, and the mode.

The Arithmetic Mean. The arithmetic mean, sometimes simply called the mean, may be defined as the sum of the measures divided by the number of measures.[3] Or it may be thought of as a point of balance that could be found if all values in the distribution were assigned the same weight and then arranged along a horizontal beam. The physicist might define it as that point in the distribution around which the moments are equal.

In the case of the sixty values previously discussed from which a frequency polygon was constructed, the mean is obtained by finding the total of the sixty measures and dividing this total by 60, thus:

$$\text{Arithmetic Mean } (A.M.) = \frac{\text{Sum of measures}}{\text{N}} = \frac{2016}{60} = 33.6$$

This is the procedure followed in computing the exact value of the arithmetic mean of any set of values. In practice, a shorter method of computation utilizing data as tabulated in a frequency distribution and yielding an approximation (rather than the exact value) of the mean is often used. This shorter method assumes that each score as tabulated in a frequency distribution has the same value as the midpoint of the c.i. in which it falls. For further convenience in calculation, the mean is first computed in c.i. units from an arbitrary base selected near the center of the distribution at the midpoint of one of the c.i.'s. The base selected is entirely arbitrary—it may be taken as any point in the distribution. We have chosen one near the center of the distribution to simplify computation.

If this method is applied to the frequency distribution in Table 18.3 the arrangement shown in Table 18.5 is obtained.

In the above tabulation the *d* column represents the number of c.i. units each c.i. is located above or below the c.i. arbitrarily chosen as the base for calculations. In the c.i. 51–53 there are two scores. This c.i. is 6 c.i. units above the arbitrary base. Thus, in computing the *A.M.* in c.i. units from the arbitrary base, these two scores would each have a value of 6, resulting in the number 12 that appears in the fourth, or *fd*, column. In like manner, there are two scores in the c.i. 48–50, and these two scores are each 5 c.i. units above the arbitrary base, resulting in the number 10 that appears in the *fd* column. All scores tabulated in c.i.'s below the arbitrary base are represented by negative values in the *fd* column. The algebraic sum of this column

---

[3] *Ibid.,* 151 ff.

TABLE 18.5

Computation of the Arithmetic Mean from a Frequency Distribution

c.i.	*f*	*d*	*fd*
51–53	2	6	12
48–50	2	5	10
45–47	3	4	12
42–44	4	3	12
39–41	6	2	12
36–38	9	1	9
33–35	9	0	0
30–32	7	−1	−7
27–29	5	−2	−10
24–26	4	−3	−12
21–23	4	−4	−16
18–20	2	−5	−10
15–17	2	−6	−12
12–14	1	−7	−7
	60	$\Sigma fd =$	−7

Formula for Computing *A.M.*

$A.M. = M^\circ + c.i.(c)$

M° = assumed mean

$c.i.$ = size of $c.i.$

$c = \frac{\Sigma fd}{N} =$ summation of deviations from assumed mean divided by N

$$A.M. = 34 + 3\left(\frac{-7}{60}\right)$$

$$= 34 - .35 = 33.65$$

(Σfd) divided by the number of cases indicates how far the computed mean will deviate from the assumed mean (base) in terms of c.i. units. From the tabulation, this deviation in c.i. units from the arbitrary base (assumed mean) is defined as:

$$\text{Deviation in c.i. units from base} = \frac{\Sigma fd}{N}$$

Carrying through this computation for the data under consideration shows that:

$$\text{Deviation in c.i. units from base} = \frac{\Sigma fd}{N} = \frac{-7}{60} = -.117$$

This is interpreted to mean that the *A.M.* is .117 of a class interval below the midpoint of the arbitrary base (see formula in illustrative problem). In order to transmute this deviation (—.117) into raw score units we would multiply it by 3 (the size of the class interval). Thus, in terms of raw score units, the deviation is —.35. The mean, as computed by this method, is therefore .35 raw score units below the midpoint of the 33–35 c.i. As the midpoint of this is 34, the mean is 34 — .35 = 33.65. This approximation does not agree exactly with the exact method in which all raw data are added and the sum is divided by the number of cases; but the approximation is sufficiently close to justify its use in many cases. The student may note, however, that essentially the same procedures are used in both solutions. The procedure in using the "exact method" may be thought of as involving the computation of a mean by assuming the mean to be zero, computing the deviations from zero in raw score units, and dividing their sum by N as in the short method.

The Median. The median is a measure of central tendency defined as that score (or value) that exceeds, and is exceeded by, half the measures; that is, it is that point in the distribution above and below which 50 per cent of the values lie. A logical (though laborious) method to determine the median consists in arranging all the raw data in rank order from lowest to highest and counting off the bottom half of the measures. The value at this point is the median. If this method is followed for the data in Table 18.1, the following arrangement of the scores is obtained:

TABLE 18.6

Arrangement of Data for the Computation of the Median Directly from Raw Data

53	45	41	38	36	34	32	29	26	21
51	44	40	38	36	34	31	29	25	20
49	43	39	37	35	33	31	28	24	19
48	42	39	37	35	33	30	28	23	16
47	42	39	37	35	33	30	27	22	15
46	41	38	36	34	32	30	26	21	13

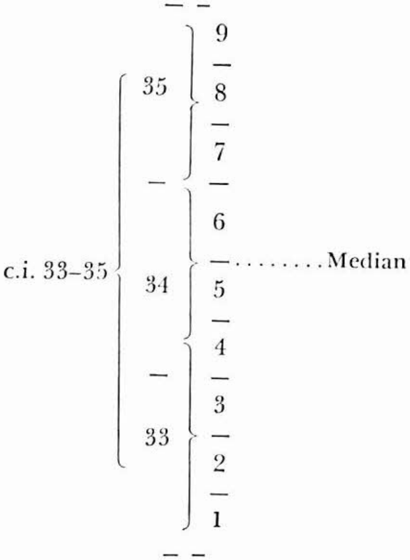
Counting from the lowest score up, we find that the thirtieth from the low end is 34, and the thirty-first from the low end is also 34. The median score would therefore be 34. If there had been a difference between the thirtieth and the thirty-first scores, the median would be the value halfway between these scores. If an odd number of cases were included in the original set of scores (as 61 instead of 60) the median would be the value of the middle score.

In practice, the median as well as the mean may be conveniently approximated from a tabulated frequency distribution. To illustrate this process we may use the same frequency distribution previously discussed. The computations are shown in Table 18.7.

Knowing that sixty cases are included in the distribution, it is necessary to find the score that separates the lower thirty from the upper thirty. The value of this score is the median. Counting up from the lower part of the distribution, we first fill the column *cumulative f*, which indicates for each c.i. the number of cases *in and below that c.i.*

It will be noted that twenty-five cases are included in or below the 30–32 c.i., and that thirty-four cases are included in or below the 33–35 c.i. The median, or point midway between the thirtieth and thirty-first case, must therefore be within the 33–35 c.i. Now if we assume that the nine cases in the 33–35 c.i. are distributed evenly throughout this interval, we must go up into this c.i. far enough to cover the lowest five of these nine cases in order to reach the median. This may be illustrated graphically as follows:

Cases in this c.i.



It will be noted that the "real limits" of the 33–35 c.i. are considered 32.5 and 35.5; that is, the c.i., extends one-half a score unit above and one-half a score unit below the tabulated values. This is necessary because otherwise there would be a whole unit between each pair of adjacent c.i.'s that does not logically fall in either c.i.

In this case, the median would be five-ninths of the size of the c.i., added to the lower limit of the c.i., or Median $= 5/9(3) + 32.5 = 34.2$. This value, 34.2, differs slightly from the value computed by arranging the scores in rank order, but the approximation is sufficiently close to justify its use in most instances. Exactly the same procedure can be followed in defining other points in the distribution, for example, the twenty-fifth percentile—the point in the distribution below which 25 per cent of the scores lie and above which 75 per cent lie.

TABLE 18.7

Computation of the Median from a Frequency Distribution

Class Intervals	*f*	*Cumulative f*
51–53	2	60
48–50	2	58
45–47	3	56
42–44	4	53
39–41	6	49
36–38	9	43
33–35	9	34
30–32	7	25
27–29	5	18
24–26	4	13
21–23	4	9
18–20	2	5
15–17	2	3
12–14	1	1
Total	60	

Formula for Median

$$Md. = l + c.i. \frac{\left(\frac{N}{2} - F\right)}{f_m}$$

$l =$ lower limit of c.i. within which median lies.

$c.i. =$ class interval

$\frac{N}{2} = \frac{1}{2}$ of the scores.

$F =$ no. of scores in all $c.i.$'s below l.

$f_m =$ no. of scores in $c.i.$ in which median falls.

$$Md. = 32.5 + 3\,\frac{(60/2 - 25)}{9}$$

$$= 32.5 + 1.7 = 34.2$$

The Mode. A third measure of central tendency is the mode, which is defined as the measure appearing most frequently. This value, as well as the mean and the median, may be determined directly from the raw data (if one value appears more often then any other) or may be approximated from a frequency distribution of the data.

In computing the mode directly from the raw data, the values are inspected to determine which one appears most frequently. Sometimes, as in the case of the values shown in Table 18.1, several of the measures appear an equal number of times. In this case, each of the values 30, 33, 34, 35, 36, 37, 38, 39 appears three times. It is incorrect, therefore, to say that any one of these is the mode. Furthermore, there is reason to believe that if a larger sample than sixty inspectors had been included in the distribution, and if the c.i. used in forming the distribution were smaller than 3, the frequency polygon obtained would be more even in curvature and only one single high point would be found. Under such circumstances, this high point would be the mode. An approximation of this value may be obtained from a frequency distribution by means of the following empirical formula:

$$\text{Mode} = 3(\text{Median}) - 2(\text{Mean})$$

In the case of the data we have been discussing, this formula gives the following value for the mode:

$$\text{Mode} = 3(34.2) - 2(33.65) = 35.30.$$

When to Use the Mean, Median, and Mode. Why is it necessary to have three different measures to indicate the central tendency of a set of data? The answer is that each is best adapted to certain uses; that is in some cases one may be most representative of a set of data, while in other cases another measure may be most suitable. The mean is ordinarily used if the distribution is approximately normal. (If the distribution is perfectly normal, the three measures of central tendency have the same value.) If, on the other hand, there is a preponderance of extreme cases at either end of the distribution, the mean may give an incorrect impression of the central tendency of the data. Under these circumstances, the median or mode is more suitable. Consider, for example, the following yearly incomes of five persons:

$800 $900 $850 $750 $5,000

The mean for these five incomes is

$$\frac{\$800 + \$900 + \$850 + \$750 + \$5{,}000}{5} = \$1{,}660.$$

This figure, though an accurate statement of the mean, is not typical of the group as a whole because it is so markedly affected by the one income of $5,000 that is considerably larger than the other four. The median income is $850, and this value is more typical for the group as a whole than is the mean income of $1,660. If a great deal of data were available for computation, it would also be enlightening to know the mode, or most common income. Certain other principles also help determine which measure of central tendency is most appropriate in any specific case. We may generalize in the above illustration by saying that if a distribution is very much *skewed* (that is, contains more cases at one extreme than at the other), the median or mode is more likely to give a representative picture of the typical score than is the mean.

MEASURES OF VARIABILITY

In addition to a measure or value to represent the central tendency of a set of data, there is also quite frequently a need for some measure of the spread, or variability, of the data. The need for a measurement of this type may be seen by comparing the data shown in Table 18.1 and tabulated in Table 18.5 (the mean of which, computed from the frequency distribution, was found to be 33.65) with another set of data that, for purposes of illustration, we might assume to consist of twenty-one scores of 33, and thirty-nine scores of 34, making sixty scores in all. The mean of sixty such scores may readily be found to be 33.65.

$$\frac{(21)(33) + (39)(34)}{60} = 33.65$$

While both distributions have the same mean, they differ markedly in variability or spread. The former distribution is made up of scores varying from 13 to 53, while the latter consists entirely of scores of 33 and 34. A quantitative measure of variability is therefore of considerable value. Statistical procedures have been designed that yield a single value descriptive of this variability; as in the cases of means and medians these measures tell us something about the group as a whole.

The Standard Deviation. The Standard Deviation is the most widely used measure of variability. It is defined as the square root of the mean square deviation. Defined by formula:

$$\text{Standard Deviation} = S.D. = \sigma = \sqrt{\frac{\Sigma D^2}{\text{N}}}$$

from their mean" and N is the number of cases. S.D. and σ are abbreviations for the Standard Deviation. They are used interchangeably.

Although the Standard Deviation may be computed directly from
where ΣD^2 is read "the sum of the squared deviation of the scores
a set of raw data by means of the formula $S.D. = \sqrt{\frac{\Sigma D^2}{\text{N}}}$, this process

is laborious. For example, in the case of the set of data we have been using for illustrative purposes (tabulated in Table 18.1), we would proceed by determining the difference between each raw score and the mean of the sixty scores, squaring these differences, summing the sixty squared differences, dividing by 60, and extracting the square root of the quotient. The first score tabulated is 15. The difference between this value and the mean of the sixty scores (as computed directly from the raw data) is $D = 33.6 - 15.0 = 18.6$. D^2 would therefore be $(18.6)^2 = 345.96$. This must be repeated for every one of the sixty scores before the sum of the squared deviations can be obtained.

Because of the excessive labor in computing the S.D. directly from the raw data, a simple process that approximates the true value of the S.D. has been developed. This is used in the computations shown in Table 18.8, in which is computed the Standard Deviation of the data shown previously in Table 18.5.

TABLE 18.8

Computation of the Standard Deviation from a Frequency Distribution[4]

c.i.	*f*	*d*	*fd*	*fd*²
51–53	2	6	12	72
48–50	2	5	10	50
45–47	3	4	12	48
42–44	4	3	12	36
39–41	6	2	12	24
36–38	9	1	9	9
33–35	9	0	0	0
30–32	7	−1	−7	7
27–29	5	−2	−10	20
24–26	4	−3	−12	36
21–23	4	−4	−16	64
18–20	2	−5	−10	50
15–17	2	−6	−12	72
12–14	1	−7	−7	49
Total = 60			$\Sigma fd = -7$	$\Sigma fd^2 = 537$

$$\text{Mean} = M^\circ + c.i.\left(\frac{\Sigma fd}{\text{N}}\right)$$

$$\text{Mean} = 34 + 3\left(\frac{-7}{60}\right) = 33.65$$

$$\begin{aligned}\text{S.D.} = \sqrt{\frac{\Sigma D^2}{\text{N}}} &= c.i.\sqrt{\frac{\Sigma fd^2}{\text{N}} - \left(\frac{\Sigma fd}{\text{N}}\right)^2}\\ &= 3\sqrt{537/60 - (.117)^2}\\ &= 3\sqrt{8.950 - .014}\\ &= 3\sqrt{8.936}\\ &= 3(2.99)\end{aligned}$$

$$\text{S.D.} = 8.97$$

---

[4] The derivation of the formula for computing the standard deviation by the method used in Table 18.8 is given below. Working in c.i. units rather than raw score units, let

$D_1, D_2, \ldots D_n$ = deviations of the scores in c.i. units from their mean.

c = difference in c.i. units between the mean of the scores and the midpoint of zero c.i.; i.e. the arbitrary base.

$d_1, d_2, \ldots d_n$ = deviations of the scores in c.i. units from the arbitrary base.

Then

$$\begin{aligned}D_1 &= d_1 - c\\ D_2 &= d_2 - c\\ &\cdots\cdots\\ D_n &= d_n - c\end{aligned}$$

If the above equations are squared on both sides, we have

$$\begin{aligned}D_1^2 &= (d_1 - c)^2 = d_1^2 - 2d_1c + c^2\\ D_2^2 &= (d_2 - c)^2 = d_2^2 - 2d_2c + c^2\\ &\cdots\cdots\\ D_n^2 &= (d_n - c)^2 = d_n^2 - 2d_nc + c^2\end{aligned}$$

Summating the above to determine the sum of the squared differences, we have

$$D_1^2 + D_2^2 + \ldots D_n^2 = (d_1^2 + d_2^2 + \ldots d_n^2) - (2d_1c + 2d_2c + \ldots 2d_nc) + \text{N}c^2$$

The above, written with summation signs, becomes:

$$\Sigma D^2 = \Sigma d^2 - 2c\Sigma d + \text{N}c^2 \quad (1)$$

The Standard Deviation is the most commonly used measure of variability. Usually when the mean value of a set of data is given, the S.D. is also given to indicate the variability of the data.

COMPARABLE SCORES

The S.D. performs another useful function—it can be used in comparing individual scores from different distributions. For example, suppose that two inspectors from departments *A* and *B*, who are working at different inspection jobs, detect respectively forty-five and eighty-nine defects during a week of work. How can we compare the efficiency of these two employees? It will be seen immediately that a

---

Now it will be remembered from Table 18.5 that *c*, which is the difference in c.i. units between the mean of the data and the arbitrary base (i.e. the mean in c.i. units away from the arbitrary base) was determined by adding the *d* values of the scores and dividing this sum by the number of scores. Since there are f_1 scores at d_1 deviation; f_2 scores at d_2 deviation; etc.; this summation is given by $f_1d_1 + f_2d_2 + \dots f_nd_n = \Sigma fd$, which, divided by N, gives $\frac{\Sigma fd}{N}$.

We may therefore substitute in (1) $\frac{\Sigma fd}{N}$ for *c*, giving

$$\Sigma D^2 = \Sigma d^2 - 2\frac{\Sigma fd}{N}\Sigma d + N\left(\frac{\Sigma fd}{N}\right)^2$$

But the Σd and Σfd are the same, since Σfd is only a simpler method of determining the Σd that involves grouping together all scores of the same *d*, and multiplying this *d* by *f*, the number of such scores. For the same reason, Σd^2 is the same as Σfd^2. The above equation therefore may be written:

$$\Sigma D^2 = \Sigma fd^2 - 2\frac{\Sigma fd\,\Sigma fd}{N} + N\left(\frac{\Sigma fd}{N}\right)^2$$

$$\Sigma D^2 = \Sigma fd^2 - 2\frac{(\Sigma fd)^2}{N} + \frac{(\Sigma fd)^2}{N}$$

If both sides of the above equation are divided by N, we have

$$\frac{\Sigma D^2}{N} = \frac{\Sigma fd^2}{N} - 2\left(\frac{\Sigma fd}{N}\right)^2 + \left(\frac{\Sigma fd}{N}\right)^2$$

$$\frac{\Sigma D^2}{N} = \frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2$$

Extracting the square root of both sides

$$S.D. = \sqrt{\frac{\Sigma D^2}{N}} = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}$$

which is the equation used in the computations accompanying Table 18.8 to determine the standard deviation of the distribution in c.i. units. The value yielded by this expression is then multiplied by the size of the c.i. to give the *S.D.* in raw score units.

direct comparison of the figures 45 and 89 is not valid, because the two inspection jobs may be very different. It will also be seen that we can say little concerning the position of these inspectors in their respective groups without knowing their relation to the mean of their group in inspection work. To make a comparison, then, we must first compute the mean number of defects spotted by all inspectors in Department *A*, and the mean number spotted by all inspectors in Department *B*. Suppose that those means are respectively 38 and 95. We thus see that the inspector from Department *A* is $45 - 38 = 7$ pieces *above* the mean for that department, and that the inspector from Department *B* is $89 - 95 = -6$, or 6 pieces *below* the mean of inspectors from that department. We can thus say, at this point, that the inspector from Department *A* is above average in ability on the job and that the inspector from Department *B* is below average. But how about their relative distance from the average? To answer this question we must compute the S.D.'s of the two distributions and determine how many S.D.'s each inspector is above or below average.

Suppose we find the S.D. of the operators in Department *A* to be 5.5 pieces. Our first inspector is therefore $\frac{45 - 38}{5.5} = 1.27$ S.D.'s above average. If the S.D. of the inspectors in Department *B* is 9.5, the inspector from the group who detected eighty-nine pieces is $\frac{89 - 95}{9.5} = -.63$ or .63 S.D.'s below average. *The deviation of a score from the mean of the distribution expressed in S.D. units results in a measurement that is comparable with similarly determined measurements from other distributions.* Thus, we may say that our first inspector is about *twice* as far above average, in terms of comparable scale units, as the second operator is below average. Scores computed in this manner are known as *z*-scores. The formula for a *z*-score is as follows:

$$z\text{-score} = \frac{\text{Raw Score} - \text{Mean of Raw Scores}}{S.D.\text{ of Raw Scores}}$$

The *z*-score is helpful not only when comparing scores from one distribution to another but also when, for any reason, it is desired to combine scores with the same or differential weighting. A typical example of an industrial situation that requires this technique is in the combination of items used in a merit-rating blank. Suppose that each employee has been rated by his supervisor on a chart containing items such as the following.

Industriousness 0	10	20	30	40	50
	Always loafs if not watched	Often loafs when not watched	Sometimes loafs and sometimes works when not watched	Usually is hard at work	Always is hard at work

Knowledge of Job 0	10	20	30	40	50
	Knows little about the job	Knows routine only	Is fairly well informed on his work	Well informed on details relative to his work	Thorough knowledge of present job and related work

We may suppose, for purposes of illustration, that it is now desired to combine these two traits into an over-all merit rating. (If more than two traits are included in the chart, as is usually the case, the procedure is identical.) Suppose that an employee, Mr. A, has received 40 points on industriousness and 30 points on knowledge of job, making a total of 70 points if the ratings are added directly. Suppose that another employee, Mr. B, has received 30 points on industriousness and 40 points on knowledge of job, which also results in a total of 70 points if added directly. It is clear that such direct and immediate combination of ratings would result in identical over-all ratings for these two employees. The question we may raise is whether such a statement of equal ratings is justified. The answer is that it is not. If the mean rating of all employees on industriousness was 33 with a S.D. of 3, then A's rating would be $\frac{40 - 33}{3} = 2.33$, or 2.33 S.D.'s above the mean and B's would be $\frac{30 - 33}{3} = -1.00$, or 1.00 S.D. below the mean on this trait. If the mean rating for all employees on knowledge of job were 25, with a S.D. of 6, A would be $\frac{30 - 25}{6} =$

+.83, or .83 S.D.'s above average in knowledge of job while B would be $\frac{40 - 25}{6} = 2.50$ or 2.50 S.D.'s above average in this respect. Now, the proper combination of the two traits, if we wish to weight them equally, would be:

Employee	Rating in Industriousness	Rating in Knowledge of Job	z-Score in Industriousness	z-Score in Knowledge of Job	Sum of Scores for Both Units
A	40	30	+2.33	+.83	+3.16
B	30	40	−1.00	+2.50	+1.50

This transfer of ratings into *z*-scores and the adding of the *z*-scores shows that the two employees A and B are not equal in rating (as we would infer if the raw ratings were added), but rather that A is definitely higher than B. The procedure described has assumed that the two trait ratings, being combined, should be given equal weight, and the procedure shows how they can be combined with equal weight into a composite score. One might think that conversion of raw scores to *z*-scores is not necessary if the raw scores are to be given equal weight in the combination score. Actually, if we do not give the raw scores equal weight by converting them into *z*-scores, the scores will weight themselves according to the size of their respective Standard Deviations. In other words, if combined directly, the raw scores will be weighted too much or too little, depending upon their position relative to the means of their respective distributions and upon the variability of the distribution of which they are a part. When scores are combined, they are *always* weighted in some manner, whether we deliberately weight them or not. It is highly important, therefore, to weight them deliberately (either with equal weight or otherwise) by converting them into *z*-scores and then combining them.

It does not follow from the above discussion that combined scores should always be weighted equally. Indeed, it is often desirable to weight various scores according to some plan that has been decided upon before the scores are combined. When this is desired, such weighting can be accomplished very easily by multiplying each *z*-score by the appropriate weight before they are combined. In our illustrative case, suppose that we have decided that *industriousness* should be given twice as much weight as *knowledge of job* in determining the total rating. This would be accomplished as follows:

Employee	z-Score in Industriousness	z-Score in Knowledge of Job	Weighted z-Score in Industriousness	Weighted z-Score in Knowledge of Job	Combined Weighted z-Scores
A	+2.33	+.83	+4.66	+.83	+5.49
B	−1.00	+2.50	−2.00	+2.50	+.50

The combined ratings so obtained show a still greater difference between employees A and B than was obtained when the scores were equally weighted. If, on the other hand, it is desired to give the rating on *knowledge of job* twice as much weight as the rating on *industriousness*, the following computations would be made:

Employee	z-Score in Industriousness	z-Score in Knowledge of Job	Weighted z-Score in Industriousness	Weighted z-Score in Knowledge of Job	Combined Weighted z-Scores
A	+2.333	+.833	+2.333	+1.666	+4.0
B	−1.000	+2.500	−1.000	+5.000	+4.0

This last procedure results in giving identical total scores to employees A and B, which, it will be remembered, also occurred when the raw scores were added directly ($30 + 40 = 70$). Now, if we remember that the assumed Standard Deviation for the *knowledge-of-job* ratings was 6, which is twice as great as the Standard Deviation of 3 assumed for the *industriousness* ratings, we can see why a direct combination of raw scores gives the equality that is obtained when the respective *z*-scores are weighted in the ratio 2:1. In this instance the original ratings whose Standard Deviation is 6 are automatically given twice as much weight (when raw scores are combined directly) as the original ratings whose Standard Deviation is only 3. Thus, we obtain the same final result (equality) by direct combination that is obtained when the *knowledge-of-job* ratings are deliberately given a weight twice as great as the *industriousness* ratings.

Many other problems arise in which it is necessary to weight scores to achieve a particular result. For example, a company faced with the problem of selecting a number of electrical apprentices desires to give this training to those boys who have the greatest aptitude for the job and who are therefore most likely to succeed. Careful consideration and discussion of the problem by management and supervision resulted in the decision that four factors should determine whether an employee should be given this training. These four factors were general intelligence, present knowledge of electricity, previous merit rating, and seniority with the company. It was further decided in conference that, although all of these four factors should be considered, they are not of equal importance. It was decided that a fair weighting of their relative importance was as follows:

General intelligence	40%
Knowledge of electricity	30%
Merit rating	20%
Seniority or service with the company	10%

To score the applicants according to this plan, each was given a general intelligence test and a test covering technical phases of electricity. Merit ratings and seniority were obtained from the company records. Each of the four scores was converted into a z-score and the four resulting z-scores were respectively multiplied by 40, 30, 20, and 10. For each employee the sum of the weighted z-scores was used in indicating whether or not he was given the apprenticeship training.

PERCENTILES

The discussion of comparable scores should have made clear the fact that a raw score on any test is relatively meaningless unless it is interpreted in terms of its location in a distribution of other scores made by other people. If a test consists of seventy-five very easy questions, a score of 65 might be near the bottom of the distribution and hence should be interpreted as a very low score. On the other hand, if a test consists of seventy-five very difficult questions, a score of 65 might be at or near the top of the distribution and should therefore be considered a very high score. In other words, a raw score of 65 might be a low score or a high score, depending upon the distribution of scores from which it is drawn.

One convenient and widely used method of interpreting a raw score is by using *percentile ranks*. A percentile rank may be defined as the number showing the percentage of the total group equal to or below the score in question. Thus, on a certain test, if 65 per cent of the total group scored 129 or below, the score of 129 would be at the 65th percentile, or would have a percentile rank of 65. The 50th percentile, it will be noted, is the same as the median as previously defined.

A convenient, practical method of determining by close approximation the percentile equivalents of a set of raw scores makes use of a cumulative frequency distribution such as the one tabulated in Table 18.9. This tabulation is based on the same distribution previously used in Table 18.3.

TABLE 18.9

Cumulative Frequency Distribution Used in Determining Percentile Ranks of Raw Scores

(1) *Class Intervals*	(2) *f*	(3) *Cumulative f*	(4) *Per Cent*
51–53	2	60	100
48–50	2	58	96
45–47	3	56	93
42–44	4	53	88
39–41	6	49	81
36–38	9	43	71
33–35	9	34	56
30–32	7	25	42
27–29	5	18	30
24–26	4	13	22
21–23	4	9	15
18–20	2	5	8
15–17	2	3	5
12–14	1	1	2
	Total = 60		

The per cent values, in column (4) of Table 18.9, are obtained by dividing each of the values in the cumulative *f* column, column (3), by the total of column (2), in this instance 60. The per cent values in column (4) are then plotted against the upper limits of the class intervals, as shown in Fig. 18.3. From Fig. 18.3 the percentile ranks of the raw scores may be read directly, with sufficient accuracy for most purposes. When this has been done, the raw scores with corresponding percentile ranks may be tabulated as illustrated in Table 18.10.

In Table 18.10, only certain percentile ranks are given, because these are usually sufficient for ordinary purposes. The intervening percentile ranks, however, may also be read from Fig. 18.3, if this is necessary.

The manual published with standardized tests usually includes percentile tables making possible the conversion of raw scores into percentile ranks. Since the percentile rank of a given raw score is dependent upon the nature of the group used in constructing the conversion table, several raw-score-to-percentile-rank conversion

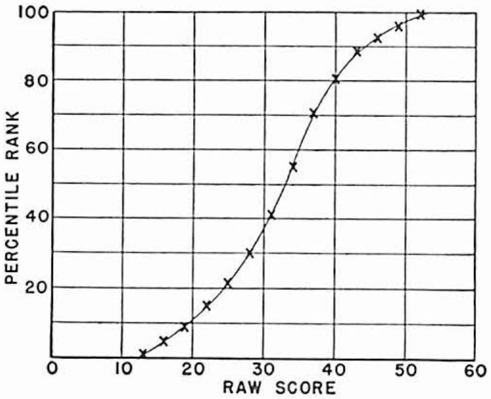


FIG. 18.3. Chart for converting raw scores into percentile ranks for illustrative data in Table 18.10.

tables, based on different groups, are often published with standardized tests.

TABLE 18.10

Raw Scores and Percentile Rank Equivalent Read from Fig. 18.3

Percentile Rank	*Raw Score*
100	53
98	52
95	48
90	44
80	40
70	37
60	35
50	33
40	31
30	28
20	24
10	20
5	17
2	14
1	13

CORRELATION

In numerous experimental situations two variable quantities are so related that they vary, or tend to vary, with each other. A common problem in industrial psychology is to reduce to a simple and meaningful statement the facts that have been discovered concerning such a functional relationship. Suppose that a number of punch-press operators during a given period of time have each punched a certain number of pieces and have each mispunched, or otherwise wasted, a certain number of pounds of stock material. In such a situation, it might be of considerable importance for management to know whether any relationship exists (and, if so, how much) between quantity of work done and amount of material wasted. Indeed, the company's policy with respect to speed of work recommended as well as the quality control in the form of penalty or bonus might well be formulated correctly only in the light of specific knowledge of the relationship between speed and accuracy.

Consider a department employing eight operators for whom the following figures for production and waste in pounds are available:

Operator	*Production*	*Waste*
1	95	3.0
2	103	4.5
3	88	3.5
4	98	4.0
5	93	3.0
6	107	4.5
7	114	4.0
8	106	5.0

It is difficult, if not impossible, to determine from a gross inspection of these two columns of figures whether any relationship exists between speed and waste. It is necessary to employ some type of graphic or computational procedure to determine the amount of relationship that may exist between these two sets of data. One simple and sometimes satisfactory method consists of a simple plotting of the values on co-ordinate axes and rough inspection of the results. If we let production be represented on the x or horizontal axis, and waste on the y or vertical axis, then the production and waste of each operator will locate him on a chart, giving the result shown in Fig. 18.4.

A plot such as is shown in Fig. 18.4 gives a much better indication of the presence or absence of a relationship between the data than can be obtained from the columns of raw data from which the chart was prepared. The chart shows that some relationship does exist, and it is even possible to draw in by inspection a line or curve that represents this relationship in an approximate form.

Although this simple method of studying the relationship between two variables is sometimes adequate for very simple problems or for those that involve only a small amount of data, it is not adequate for

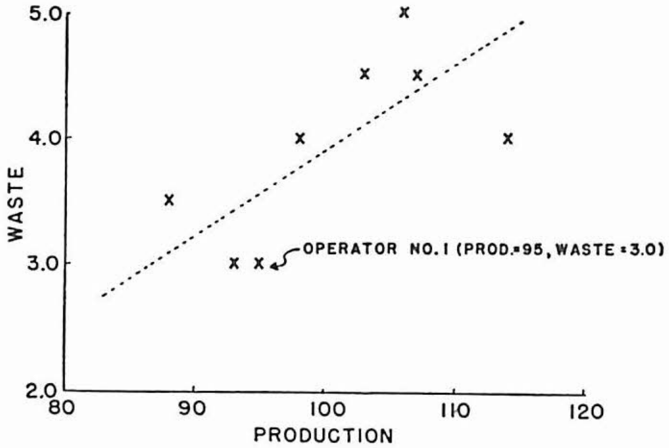


FIG. 18.4. A plot of the production and waste records for the eight punch-press operators shown above.

an exact study because it does not result in a quantitative statement of the degree of relationship. The slope of the dotted line cannot be considered such a quantitative statement because: (1) this line is drawn in by inspection, and (2) its slope depends upon the units of measurement on both the *x* and *y* axes.

Two commonly used quantitative methods for measuring the degree of relationship between two paired sets of data are rank-order correlation and product-moment correlation.

Rank-Order Correlation. The use of this method may be described by applying it to the data for the eight punch-press operators.

In Table 18.11 the two columns headed *Rank* give, respectively, the rank of the operators on the two measures (production and waste). The highest producing operator (in this case the seventh in the list) is given a rank of 1, the second highest a rank of 2, and so on. In like manner, the rank of each operator in wastage is placed in the waste-rank column. In case two or more operators are tied for a given rank (as in the case of the second and sixth operators who are tied at 4.5 pounds of waste each), the tied scores are all given the same rank, which is the average of the ranks that would have been assigned to the tied scores if they had not been tied. The values in the D^2 column are obtained by squaring each D value. The sum of the D^2 column is then determined and the correlation computed by means of the formula:

$$\text{Rho} = 1 - \frac{6\Sigma D^2}{N(N^2 - 1)}.$$

N is the number of cases entering into the computation.

This formula for the rank-order correlation is an empirical formula. It yields a value of +1.00 if the data are in exactly the same rank order. (The reason for this may be seen from the fact that if all ranks are the same, all D's are zero, all D^2 values are zero, ΣD^2 is zero, and the formula becomes $1 - 0 = 1$.) If the data are in exactly reverse order (that is, if the individual who ranks highest on one series is lowest on the other, and so on) the formula will yield a value of − 1.00, but if no relationship exists between the two sets of data, a correlation of zero will be found.

TABLE 18.11

Computation of Rank-Order Correlation Rank

Operator	*Production*	*Waste*	*Rank in Production*	*Rank in Waste*	*Difference in Rank (D)*	$(D)^2$
1.	95	3.0	6	7.5	1.5	2.25
2.	103	4.5	4	2.5	1.5	2.25
3.	88	3.5	8	6.0	2.0	4.00
4.	98	4.0	5	4.5	.5	.25
5.	93	3.0	7	7.5	.5	.25
6.	107	4.5	2	2.5	.5	.25
7.	114	4.0	1	4.5	3.5	12.25
8.	106	5.0	3	1.0	2.0	4.00
						25.50

$$\text{Rho} = 1 - \frac{6\Sigma D^2}{N\,(N^2 - 1)} = 1 - \frac{153}{504} = .70$$

The use of this formula is ordinarily more satisfactory than a simple plotting of one variable against the other because it yields a quantitative statement of the degree of relationship and not simply a graphic representation that cannot be reduced to a numerical statement.

If an appreciable number of cases are involved, however, the rank-order method of computing the degree of relationship is extremely laborious. For this reason—and for other reasons of a mathematical nature—it is ordinarily used only when the data are limited to a very few cases (less than thirty).

The Product-Moment Coefficient. This is the most widely used measure of relationship. Like the rank-order correlation, it may vary from +1.00 (indicating perfect positive relationship) through zero (indi-

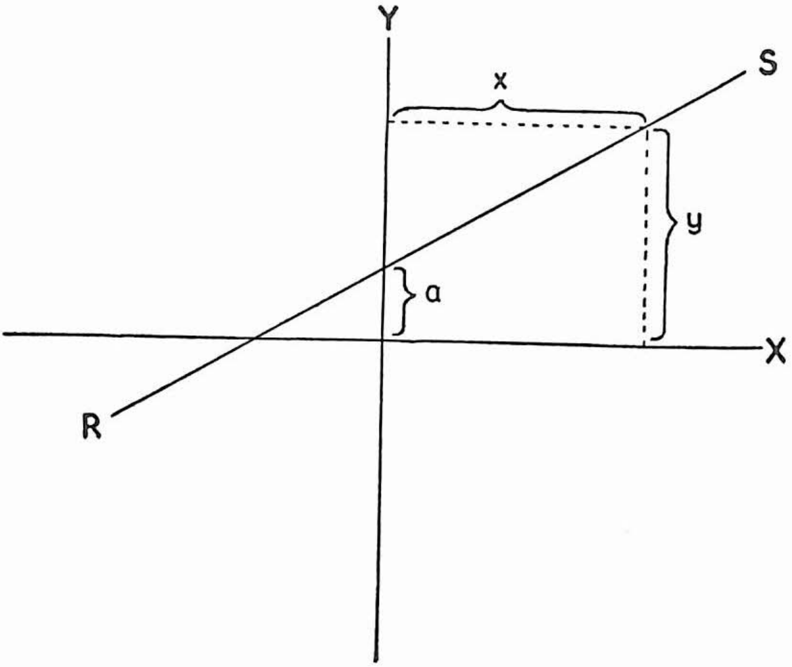


FIG. 18.5. The slope of the line RS is defined as $b = \frac{y - a}{x}$

cating no relationship) to —1.00 (indicating perfect negative relationship). The product-moment correlation, represented by the symbol r, may be defined in several ways. One of the simplest definitions is that *r is the slope of the straight line that best fits the data after the data have been plotted as z-scores on co-ordinate axes;* that is, it is the tangent of the angle made by this line with the base line.

Several terms in this definition require further definition. By *slope* is meant steepness with which the line rises. The slope of a straight line drawn in any manner across co-ordinate paper is defined as the distance, y, from any given point on the line to the x intercept, minus the distance, a, from the origin to the y intercept, divided by the distance, x, from the point on the line to y intercept. Thus the slope, which we will call b, is defined in Fig. 18.5 as follows:

$$b = \frac{y - a}{x}$$

It should be remembered that, on co-ordinate axes, distances above and to the right of the origin are positive, while distances measured below and/or to the left of the origin are negative. The slope of any line that *rises* as it goes from left to right will therefore be positive (the greater the rise in a given distance to the right the larger the positive value of the slope) and the slope of any line that *falls* as it goes from left to right will be negative (the greater the fall in a given distance to the right, the greater the negative value of the slope).

TABLE 18.12

Production and Waste for Eight Punch-Press Operators, with Corresponding z-Scores of the Production and Waste Figures

Operator	*Production*	*Waste*	*z-Score in Production*	*z-Score in Waste*
1.	95	3.0	−.69	−1.38
2.	103	4.5	+.31	+.82
3.	88	3.5	−1.56	−.65
4.	98	4.0	−.31	+.09
5.	93	3.0	−.94	−1.38
6.	107	4.5	+.81	+.82
7.	114	4.0	+1.69	+.09
8.	106	5.0	+.69	+1.56
Mean	100.5	3.94		
S.D.	8.0	.68		

By line of *best fit* in the definition is meant a line so drawn that the sum of the squared deviations in a vertical direction from the original points to the line is less than the sum would be for any other straight line that might be drawn.

A rough approximation of the value of r may be obtained by plotting the z-scores of the two variables, fitting a straight line to these points by inspection, and graphically measuring the slope of this straight line. Although this method is never used in practical computation (because it is both inaccurate and laborious), the application of it to a set of representative data may serve to clarify the meaning of the correlation coefficient, r. Returning to the data for which we have previously computed the rank-order correlation (see Table 18.11 on p. 537), we first compute the z-scores for each measure as shown in Table 18.12.

These pairs of z-scores are used as the x and y values for eight points that are plotted on co-ordinate axes as in Fig. 18.6. The straight line that seems best to fit these points is then determined (as with a stretched string that is moved about until the desired location is obtained) and drawn on the graph. The correlation, r, as determined by this crude method, is obtained by measuring the slope of

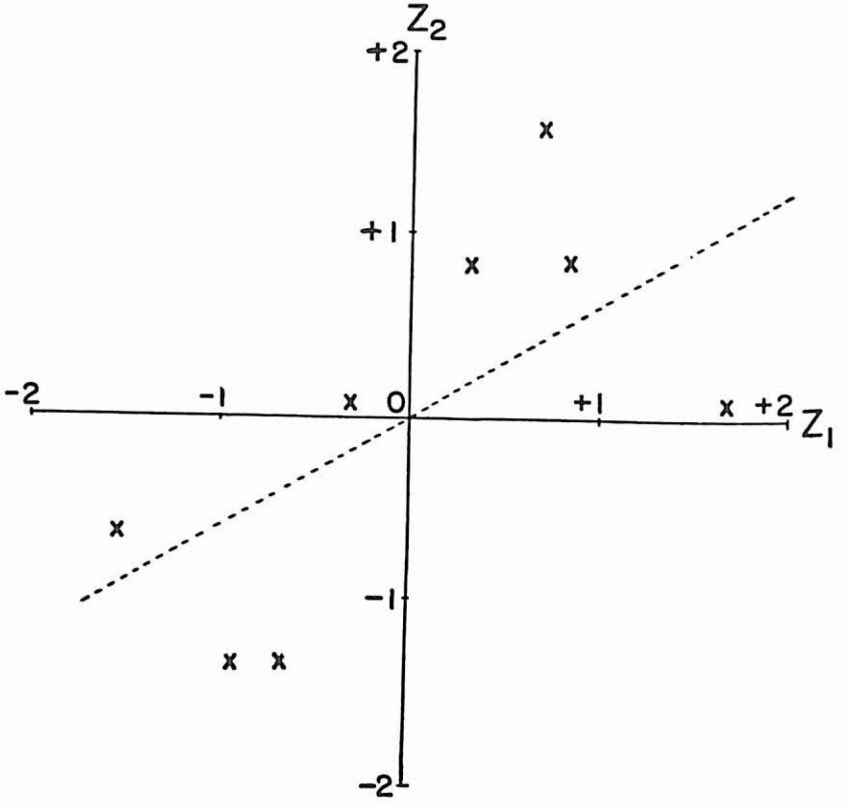


FIG. 18.6. A plot of the z-scores for production and waste records of the eight punch-press operators shown above.

this line. The procedure applied to Fig. 18.6 gives a value of $r = .61$, but it should be emphasized that this value is affected by:

1. The accuracy with which the straight line has been located, and
2. The accuracy with which the slope of the line has been measured after it has been drawn.

Points (1) and (2) both operate to eliminate the possibility of complete accuracy in this method of determining a correlation coefficient. Therefore, a mathematical method has been devised to make the computation, so that no plotting of points or graphic measurements are required. This method involves determining the equation of the straight line that, if plotted, would best fit the points, and computing the slope of this straight line of best fit from the equation.

TABLE 18.13

Computation of r by z-Score Method between Production and Waste Figures

Operator	Production	Waste	z-Score in in Production	z-Score in Waste	$(z_1 z_2)$
1.	95	3.0	−.69	−1.38	+.95
2.	103	4.5	+.31	+.82	+.25
3.	88	3.5	−1.56	−.65	+1.01
4.	98	4.0	−.31	+.09	−.03
5.	93	3.0	−.94	−1.38	+1.30
6.	107	4.5	+.81	+.82	+.66
7.	114	4.0	+1.69	+.09	+.15
8.	106	5.0	+.69	+1.56	+1.08
					+5.38

$$r = \frac{Z_x Z_y}{N} = \frac{+5.38}{8} = .67$$

It may be proved mathematically that the slope of the straight line of best fit is given by the following equation[5]: Slope $= r = \frac{\Sigma Z_x Z_y}{N}$,

---

[5] The proof of this formula is as follows:

$Z_{x1}, Z_{x2}, Z_{x3}, \ldots \quad Z_{xn}$ are the z-scores for the x variable

$Z_{y1}, Z_{y2}, Z_{y3}, \ldots \quad Z_{yn}$ are the z-scores for the y variable

The equation of any straight line is: $y = a + bx$. The problem is to find the values of the constants a and b in this equation that will give the straight line that "best fits" the data according to the criterion of best fit stated on p. 539, that is, the straight line that will give a minimum value to the sum of the squared deviations between the line and the original points.

The first point (whose co-ordinate points are Z_{x1} and Z_{y1} will deviate from the line by an amount:

$$d_1 = Z_{y1} - a - bZ_{x1}$$

this deviation squared will be

$$d_1^2 = (Z_{y1} - a - bZ_{x1})^2$$

If the sum of all such squared deviations of the points from the line be represented by u, then

$$u = \Sigma d^2 = \Sigma (Z_y - a - bZ_x)^2$$

For the values of a and b that result in u being a minimum, the derivatives of u with respect to a and b, respectively, must be zero. To find the correlation coefficient it is therefore necessary only to differentiate the equation with respect to a and to b, to set the resulting derivatives equal to zero, and to solve for b (which is the slope of the straight line of best fit). This is done as follows:

where $\Sigma Z_x Z_y$ is read "the sum of the products of the z-scores for the pairs of points or values."

Applying the formula to the data in Table 18.11, we may compute the correlation as in Table 18.13.

The value of r thus obtained by computation, .67, differs from the value of .61 obtained by plotting and inspection. The plotting and inspection method yielded a value that was somewhat in error for the data in question.

$$\frac{\partial u}{\partial a} = 0 = 2\,\Sigma(Z_y - a - bZ_x)\,(-1) \tag{1}$$

$$\frac{\partial u}{\partial b} = 0 = 2\,\Sigma(Z_y - a - bZ_x)\,(-Z_x) \tag{2}$$

(1) above becomes:

$$\begin{aligned} 0 &= -\Sigma Z_y + Na + b\Sigma Z_x \\ \Sigma Z_y &= Na + b\Sigma Z_x \end{aligned} \tag{3}$$

(2) above becomes:

$$\begin{aligned} 0 &= -\Sigma Z_x Z_y + a\Sigma Z_x + b\Sigma Z_x{}^2 \\ \Sigma Z_x Z_y &= a\Sigma Z_x + b\Sigma Z_x{}^2 \end{aligned} \tag{4}$$

It will be remembered that a z-score is obtained as follows (see p. 528):

$$Z_x = \frac{X - M_x}{\sigma_x}$$

Where X is a given raw score, M_x is the mean raw score, and σ_x the standard deviation of the raw scores.

The sum of all z-scores is therefore:

$$\begin{aligned} \Sigma Z_x &= \frac{\Sigma(X - M_x)}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{NM_x}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{\dfrac{N\Sigma X}{N}}{\sigma_x} \\ &= \frac{\Sigma X}{\sigma_x} - \frac{\Sigma X}{\sigma_x} \\ &= 0 \end{aligned}$$

In a similar way it can be shown that

$$\Sigma Z_y = 0$$

By substituting 0 for ΣZ_x and ΣZ_y in (3) we find immediately that a in the equation of the straight line of best fit is zero.

While the z-score method of computing a correlation coefficient illustrated in Table 18.13 may be used with any number of cases and will yield the correct mathematical value of r, the use of this method

---

Working with (4), we find the value of b (which is the slope or the correlation coefficient) as follows:

$$\Sigma Z_x Z_y = a\Sigma Z_x + b\Sigma Z_x^2$$

since $\Sigma Z_x = 0$, this becomes:

$$\Sigma Z_x Z_y = b\Sigma Z_x^2$$

$$b = \frac{\Sigma Z_x Z_y}{\Sigma Z_x^2} \tag{5}$$

It may be shown as follows that $\Sigma Z_x^2 = N$

$$Z_x = \frac{X - M_x}{\sigma_x}$$

$$Z_x^2 = \frac{(X - M_x)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma (X - M_x)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma (X^2 - 2XM_x + M_x^2)}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma X^2 - 2M_x \Sigma X + NM_x^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\Sigma X^2 - 2\,\frac{\Sigma X}{N}\,\Sigma X + N\left(\frac{\Sigma X}{N}\right)^2}{\sigma_x^2}$$

$$\Sigma Z_x^2 = \frac{\dfrac{N\Sigma X^2 - 2(\Sigma X)^2 + (\Sigma X)^2}{N}}{\dfrac{\Sigma X^2}{N} - \left(\dfrac{\Sigma X}{N}\right)^2}$$

$$\Sigma Z_x^2 = \frac{\dfrac{N\Sigma X^2 - (\Sigma X)^2}{N}}{\dfrac{N\Sigma X^2 - (\Sigma X)^2}{N^2}}$$

$$\Sigma Z_x^2 = \frac{N^2(\Sigma X^2) - N(\Sigma X)^2}{N\Sigma X^2 - (\Sigma X)^2}$$

$$\Sigma Z_x^2 = \frac{N[N\Sigma X^2 - (\Sigma X)^2]}{[N\Sigma X^2 - (\Sigma X)^2]} \tag{6}$$

$$\Sigma Z_x^2 = N$$

Substituting the value of $\Sigma (Z_x)^2$ given in (6) in equation (5), we have the slope or $r = b = \frac{\Sigma Z_x Z_y}{N}$.

when many pairs of data are to be correlated is very laborious. It is therefore recommended, under such circumstances, that a modification of the fundamental formula $r = \frac{\Sigma Z_x Z_y}{N}$, which makes it possible to compute r from raw score values rather than z-score values, be used. One convenient formula[6] for determining the coefficient of correlation directly from the raw data is:

$$r = \frac{N\Sigma XY - \Sigma X \Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

When we apply this formula, for illustrative purposes, to the data tabulated in Table 18.13, the computations shown in Table 18.14 result:

---

[6] The proof of this formula is as follows:

$$r = \frac{\Sigma Z_x Z_y}{N}$$

$$= \frac{1}{N}\sum \frac{(X - M_x)}{\sigma_x}\frac{(Y - M_y)}{\sigma_y}$$

$$= \frac{1}{N}\sum \frac{(XY - XM_y - YM_x + M_x M_y)}{\sigma_x \sigma_y}$$

$$= \frac{1}{N}\sum \frac{\left(XY - \frac{X\Sigma Y}{N} - Y\frac{\Sigma X}{N} + \frac{\Sigma X \Sigma Y}{N^2}\right)}{\sigma_x \sigma_y}$$

$$= \frac{1}{N}\frac{\Sigma XY - \frac{\Sigma X \Sigma Y}{N} - \frac{\Sigma X \Sigma Y}{N} + \frac{N\Sigma X \Sigma Y}{N^2}}{\sqrt{\frac{\Sigma X^2}{N} - \left(\frac{\Sigma X}{N}\right)^2}\sqrt{\frac{\Sigma Y^2}{N} - \left(\frac{\Sigma Y}{N}\right)^2}}$$

$$= \frac{1}{N}\frac{\frac{N\Sigma XY - \Sigma X \Sigma Y}{N}}{\sqrt{\frac{N\Sigma X^2 - (\Sigma X)^2}{N^2}}\sqrt{\frac{N\Sigma Y^2 - (\Sigma Y)^2}{N^2}}}$$

$$= \frac{1}{N}\frac{\frac{N\Sigma XY - \Sigma X \Sigma Y}{N}}{\frac{\sqrt{N\Sigma X^2 - (\Sigma X)^2}}{N}\frac{\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}{N}}$$

TABLE 18.14

Computation of r Directly from Raw Data

Operator	*Production* (X)	*Waste* (Y)	X^2	Y^2	XY
1. ..	95	3.0	9025	9.00	285.0
2. ..	103	4.5	10609	20.25	463.5
3. ..	88	3.5	7744	12.25	308.0
4. ..	98	4.0	9604	16.00	392.0
5. ..	93	3.0	8649	9.00	279.0
6. ..	107	4.5	11449	20.25	481.5
7. ..	114	4.0	12996	16.00	456.0
8. ..	106	5.0	11236	25.00	530.0
	$\Sigma X = 804$	$\Sigma Y = 31.5$	$\Sigma X^2 = 81312$	$\Sigma Y^2 = 127.75$	$\Sigma XY = 3195.0$

$$r = \frac{N\Sigma XY - \Sigma X\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\ \sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$r = \frac{8(3195) - (804)(31.5)}{\sqrt{8(81312) - (804)^2}\ \sqrt{8(127.75) - (31.5)^2}}$$

$$r = .67$$

When a considerable number of pairs of data are to be correlated, the use of a chart will still further simplify the computations. Several forms of such a chart have been prepared. One convenient form is shown in Fig. 18.7. This chart shows the computation of the correlation between time used in inspecting 300 pieces of material and the number of defective pieces detected. In using this chart the following steps should be followed:

1. Decide upon appropriate class intervals for one of the variables (using the rules given on p. 515) and write these in on either the x or the y axis.

2. Decide upon appropriate class intervals for the other variables and write these in on the axis not used in (1) above.

3. Place one tally mark on the scattergram for each pair of values being correlated. For example, if an inspector spotted thirty-three defects in 16.5 minutes, the tally mark would go in the cell that is found at the intersection of the *row* containing thirty-three defects and the *column* containing 16.5 minutes.

$$= \frac{1}{N}\ \frac{N^2\,(N\Sigma XY - \Sigma X\Sigma Y)}{N\ \sqrt{N\Sigma X^2 - (\Sigma X)^2}\ \sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{N\Sigma XY - \Sigma X\Sigma Y}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\ \sqrt{N\Sigma Y^2 - (\Sigma Y)^2}}$$

4. After all tally marks have been placed on the chart, the rows should be added horizontally and the sum of the tally marks in each row written opposite this row in the f_y column (Column *1*).

5. The tally marks in each column should be added and the sum written at the bottom of each column in the row (*A*) the f_x row.

6. The f_y column should be added and the sum written opposite N at the bottom of this column. *The value of* N *thus obtained may be checked by adding the values in the* f_x *row. The sum of these values should also give the value of* N.

7. Each value of f_y in the column so headed should be multiplied by the value of d_y opposite it, and the resultant product written in Column *3*, headed $f_y d_y$. *The sum of Column 3 is the value of* ΣY *which is used in the formula.*

8. Each value in Column *3*, the $f_y d_y$ column, should be multiplied by the corresponding value in Column *2*, the d_y column, resulting in the values for Column *4*, or the $f_y d_y^2$ column. The *sum of Column 4 is the value of* ΣY^2 *which is used in the formula.*

9. The values going into Column *5*, the fd_x column, are determined by finding, for each row, the sum of the products of the number of cases in each cell times the *x* value of that cell. For example, in the first row in which a tally mark appears, there is only a single case, which appears in the cell under an *x* value of 13. The value to go into the blank in Column *5* is therefore (1) (13) = 13. In the next row no tally marks appear; therefore this row is blank. In the next row, one tally mark appears in the cell under an *x* value of 11, two in the cell with an *x* value of 12, and one in the cell with an *x* value of 13. The value to go into the blank cell in Column *5* is therefore (1) (11) + (2) (12) + (1) (13) = 48. The remaining cells in Column *5* are filled in a similar manner.

10. The cells in Column *6*, the $fd_x d_y$ column, are filled with values obtained by multiplying each value in Column *2*, the d_y column, by the value in that same row appearing in Column *5*, the fd_x column. The value in the first cell in Column *6* is therefore (15) (13) = 195. *The sum of Column 6 is the value of* ΣXY *which is used in the formula.*

11. The values in row (*C*) are obtained by multiplying each value in row (*A*), the f_x row, by the value directly below it in row (*B*), the d_x row. The values appearing in row *A* have already been obtained (see Step *5* above). The resultant values are entered in row (*C*), the $f_x d_x$ row. *The sum of the values appearing in row (C) is the value of* ΣX *which is used in the formula.*

12. Each value in row (*B*), the d_x row, should be multiplied by

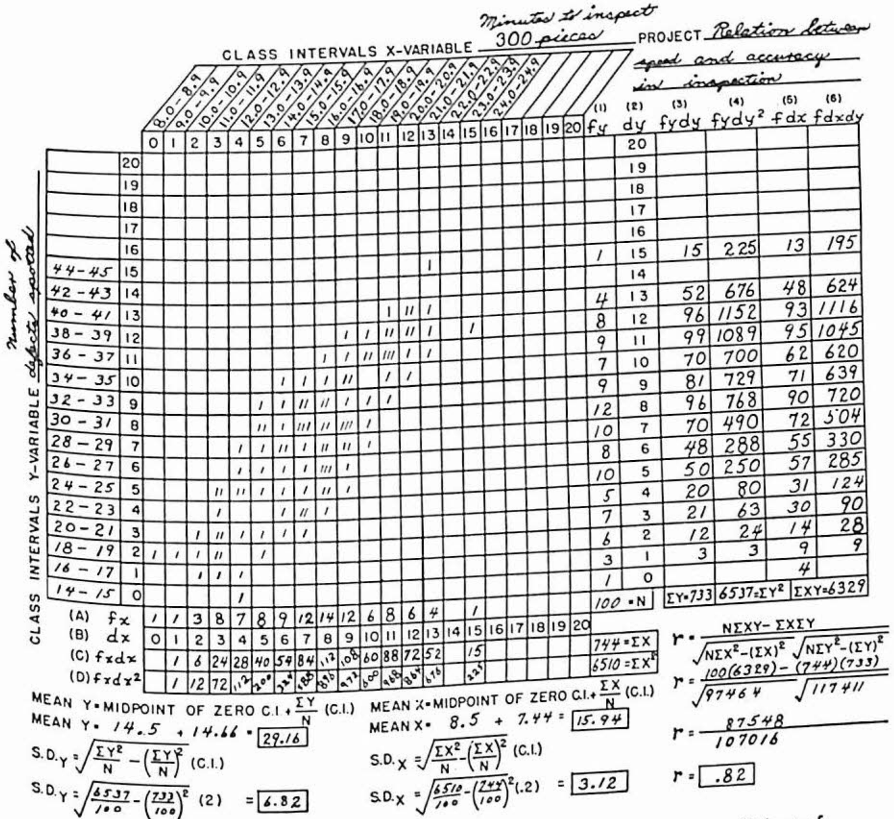

CLASS INTERVALS X-VARIABLE
Minutes to inspect 300 pieces
PROJECT Relation between speed and accuracy in inspection
CLASS INTERVALS Y-VARIABLE
Number of defects spotted
(1) fy (2) dy (3) fydy (4) fydy² (5) fdx (6) fdxdy
(A) fx
(B) dx
(C) fxdx
(D) fxdx²
100 = N
ΣY=733
6537=ΣY²
ΣXY=6329
744 = ΣX
6510 = ΣX²
MEAN Y = MIDPOINT OF ZERO C.I. + ΣY/N (C.I.)
MEAN Y = 14.5 + 14.66 = 29.16
S.D.Y = √(ΣY²/N − (ΣY/N)²) (C.I.)
S.D.Y = √(6537/100 − (733/100)²) (2) = 6.82
MEAN X = MIDPOINT OF ZERO C.I. + ΣX/N (C.I.)
MEAN X = 8.5 + 7.44 = 15.94
S.D.X = √(ΣX²/N − (ΣX/N)²) (C.I.)
S.D.X = √(6510/100 − (744/100)²) (.2) = 3.12
r = (NΣXY − ΣXΣY) / (√(NΣX² − (ΣX)²) √(NΣY² − (ΣY)²))
r = (100(6329) − (744)(733)) / (√97464 √117411)
r = 87548 / 107016
r = .82

FIG. 18.7. A chart used in the computation of a product moment coefficient of correlation.

the value directly below in row (*C*), the $f_x d_x$ row. The resultant values should be entered in row (*D*), the $f_x d_x^2$ row. *The sum of the values in row (D) is the value of* ΣX^2 *which is used in the formula.*

13. The values for N (see Step *6*), ΣY (see Step *7*), ΣY^2 (see Step *8*), ΣXY (see Step *10*), ΣX (see Step *11*), and ΣX^2 (see Step *12*) are now entered in the formula. The indicated arithmetic computations are then performed, yielding the value of *r*.

The use of this method assumes that each measure has the value of the midpoint of the class interval in which it falls. The computations indicated on the chart result in obtaining not only the value for *r* but also the Mean and the Standard Deviations of both the X and Y arrays. It will be noted that these are the same formulas previously considered on pp. 520 and 526.

APPENDIX B

TAYLOR-RUSSELL TABLES*

Tables of the Proportion Who Will Be Satisfactory Among Those Selected for Given Values of the Proportion of Present Employees Considered Satisfactory, the Selection Ratio, and r

Proportion of Employees Considered Satisfactory = .05
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05	.05
.05	.06	.06	.06	.06	.06	.05	.05	.05	.05	.05	.05
.10	.07	.07	.07	.06	.06	.06	.06	.05	.05	.05	.05
.15	.09	.08	.07	.07	.07	.06	.06	.06	.05	.05	.05
.20	.11	.09	.08	.08	.07	.07	.06	.06	.06	.05	.05
.25	.12	.11	.09	.08	.08	.07	.07	.06	.06	.05	.05
.30	.14	.12	.10	.09	.08	.07	.07	.06	.06	.05	.05
.35	.17	.14	.11	.10	.09	.08	.07	.06	.06	.05	.05
.40	.19	.16	.12	.10	.09	.08	.07	.07	.06	.05	.05
.45	.22	.17	.13	.11	.10	.08	.08	.07	.06	.06	.05
.50	.24	.19	.15	.12	.10	.09	.08	.07	.06	.06	.05
.55	.28	.22	.16	.13	.11	.09	.08	.07	.06	.06	.05
.60	.31	.24	.17	.13	.11	.09	.08	.07	.06	.06	.05
.65	.35	.26	.18	.14	.11	.10	.08	.07	.06	.06	.05
.70	.39	.29	.20	.15	.12	.10	.08	.07	.06	.06	.05
.75	.44	.32	.21	.15	.12	.10	.08	.07	.06	.06	.05
.80	.50	.35	.22	.16	.12	.10	.08	.07	.06	.06	.05
.85	.56	.39	.23	.16	.12	.10	.08	.07	.06	.06	.05
90	.64	.43	.24	.17	.13	.10	.08	.07	.06	.06	.05
.95	.73	.47	.25	.17	.13	.10	.08	.07	.06	.06	.05
1.00	1.00	.50	.25	.17	.13	.10	.08	.07	.06	.06	.05

° These tables are reproduced by permission from H. C. Taylor and J. T. Russell, "The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection: Discussion and Tables," *Journal of Applied Psychology*, 1939, **23**, 565–578.

Proportion of Employees Considered Satisfactory = .10

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10	.10
.05	.12	.12	.11	.11	.11	.11	.11	.10	.10	.10	.10
.10	.14	.13	.13	.12	.12	.11	.11	.11	.11	.10	.10
.15	.16	.15	.14	.13	.13	.12	.12	.11	.11	.10	.10
.20	.19	.17	.15	.14	.14	.13	.12	.12	.11	.11	.10
.25	.22	.19	.17	.16	.14	.13	.13	.12	.11	.11	.10
.30	.25	.22	.19	.17	.15	.14	.13	.12	.12	.11	.10
.35	.28	.24	.20	.18	.16	.15	.14	.13	.12	.11	.10
.40	.31	.27	.22	.19	.17	.16	.14	.13	.12	.11	.10
.45	.35	.29	.24	.20	.18	.16	.15	.13	.12	.11	.10
.50	.39	.32	.26	.22	.19	.17	.15	.13	.12	.11	.11
.55	.43	.36	.28	.23	.20	.17	.15	.14	.12	.11	.11
.60	.48	.39	.30	.25	.21	.18	.16	.14	.12	.11	.11
.65	.53	.43	.32	.26	.22	.18	.16	.14	.12	.11	.11
.70	.58	47	.35	.27	.22	.19	.16	.14	.12	.11	.11
.75	.64	.51	.37	.29	.23	.19	.16	.14	.12	.11	.11
.80	.71	.56	.40	.30	.24	.20	.17	.14	.12	.11	.11
.85	.78	.62	.43	.31	.25	.20	.17	.14	.12	.11	.11
.90	.86	.69	.46	.33	.25	.20	.17	.14	.12	.11	.11
.95	.95	.78	.49	.33	.25	.20	.17	.14	.12	.11	.11
1.00	1.00	1.00	.50	.33	.25	.20	.17	.14	.13	.11	.11

Proportion of Employees Considered Satisfactory = .20

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.20	.20	.20	.20	.20	.20	.20	.20	.20	.26	.20
.05	.23	.23	.22	.22	.21	.21	.21	.21	.20	.20	.20
.10	.26	.25	.24	.23	.23	.22	.22	.21	.21	.21	.20
.15	.30	.28	.26	.25	.24	.23	.23	.22	.21	.21	.20
.20	.33	.31	.28	.27	.26	.25	.24	.23	.22	.21	.21
.25	.37	.34	.31	.29	.27	.26	.24	.23	.22	.21	.21
.30	.41	.37	.33	.30	.28	.27	.25	.24	.23	.21	.21
.35	.45	.41	.36	.32	.30	.28	.26	.24	.23	.22	.21
.40	.49	.44	.38	.34	.31	.29	.27	.25	.23	.22	.21
.45	.54	.48	.41	.36	.33	.30	.28	.26	.24	.22	.21
.50	.59	.52	.44	.38	.35	.31	.29	.26	.24	.22	.21
.55	.63	.56	.47	.41	.36	.32	.29	.27	.24	.22	.21
.60	.68	.60	.50	.43	.38	.34	.30	.27	.24	.22	.21
.65	.73	.64	.53	.45	.39	.35	.31	.27	.25	.22	.21
.70	.79	.69	.56	.48	.41	.36	.31	.28	.25	.22	.21
.75	.84	.74	.60	.50	.43	.37	.32	.28	.25	.22	.21
.80	.89	.79	.64	.53	.45	.38	.33	.28	.25	.22	.21
.85	.94	.85	.69	.56	.47	.39	.33	.28	.25	.22	.21
.90	.98	.91	.75	.60	.48	.40	.33	.29	.25	.22	.21
.95	1.00	.97	.82	.64	.50	.40	.33	.29	.25	.22	.21
1.00	1.00	1.00	1.00	.67	.50	.40	.33	.29	.25	.22	.21

Proportion of Employees Considered Satisfactory = .30
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30	.30
.05	.34	.33	.33	.32	.32	.31	.31	.31	.31	.30	.30
.10	.38	.36	.35	.34	.33	.33	.32	.32	.31	.31	.30
.15	.42	.40	.38	.36	.35	.34	.33	.33	.32	.31	.31
.20	.46	.43	.40	.38	.37	.36	.34	.33	.32	.31	.31
.25	.50	.47	.43	.41	.39	.37	.36	.34	.33	.32	.31
.30	.54	.50	.46	.43	.40	.38	.37	.35	.33	.32	.31
.35	.58	.54	.49	.45	.42	.40	.38	.36	.34	.32	.31
.40	.63	.58	.51	.47	.44	.41	.39	.37	.34	32	.31
.45	.67	.61	.55	.50	.46	.43	.40	.37	.35	.32	.31
.50	.72	.65	.58	.52	.48	.44	.41	.38	.35	.33	.31
.55	.76	.69	.61	.55	.50	.46	.42	.39	.36	.33	.31
.60	.81	.74	.64	.58	.52	.47	.43	.40	.36	.33	.31
.65	.85	.78	.68	.60	.54	.49	.44	.40	.37	.33	.32
.70	.89	.82	.72	.63	.57	.51	.46	.41	.37	.33	.32
.75	.93	.86	.76	.67	.59	.52	.47	.42	.37	.33	.32
.80	.96	.90	.80	.70	.62	.54	.48	.42	.37	.33	.32
.85	.99	.94	.85	.74	.65	.56	.49	.43	.37	.33	.32
.90	1.00	.98	90	.79	.68	.58	.49	.43	.37	.33	.32
.95	1 00	1.00	.96	.85	.72	.60	.50	.43	.37	.33	.32
1.00	1.00	1.00	1.00	1.00	.75	.60	.50	.43	.38	.33	.32

Proportion of Employees Considered Satisfactory = .40
Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40	.40
.05	.44	.43	.43	.42	.42	.42	.41	.41	.41	.40	.40
.10	.48	.47	.46	.45	.44	.43	.42	.42	.41	.41	.40
.15	.52	.50	.48	.47	.46	.45	.44	.43	.42	.41	.41
.20	.57	.54	.51	.49	.48	.46	.45	.44	.43	.41	.41
.25	.61	.58	.54	.51	.49	.48	.46	.45	.43	.42	.41
.30	.65	.61	.57	.54	.51	.49	.47	.46	.44	.42	.41
.35	.69	.65	.60	.56	.53	.51	.49	.47	.45	.42	.41
.40	.73	.69	.63	.59	.56	.53	.50	.48	.45	.43	.41
.45	.77	.72	.66	.61	.58	.54	.51	.49	.46	.43	.42
.50	.81	.76	.69	.64	.60	.56	.53	.49	.46	.43	.42
.55	.85	.79	.72	.67	.62	.58	.54	.50	.47	.44	.42
.60	.89	.83	.75	.69	.64	.60	.55	.51	.48	.44	.42
.65	.92	.87	.79	.72	.67	.62	.57	.52	.48	.44	.42
.70	.95	.90	.82	.76	.69	.64	.58	.53	.49	.44	.42
.75	.97	.93	.86	.79	.72	.66	.60	.54	.49	.44	.42
.80	.99	.96	.89	.82	.75	.68	.61	.55	.49	.44	.42
.85	1.00	.98	.93	.86	.79	.71	.63	.56	.50	.44	.42
.90	1.00	1.00	.97	.91	.82	.74	.65	.57	.50	.44	.42
.95	1.00	1.00	.99	.96	.87	.77	.66	.57	.50	.44	42
1.00	1.00	1.00	1.00	1.00	1.00	.80	.67	.57	.50	.44	.42

Proportion of Employees Considered Satisfactory = .50

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50	.50
.05	.54	.54	.53	.52	.52	.52	.51	.51	.51	.50	.50
.10	.58	.57	.56	.55	.54	.53	.53	.52	.51	.51	.50
.15	.63	.61	.58	.57	.56	.55	.54	.53	.52	.51	.51
.20	.67	.64	.61	.59	.58	.56	.55	.54	.53	.52	.51
.25	.70	.67	.64	.62	.60	.58	.56	.55	.54	.52	.51
.30	.74	.71	.67	.64	.62	.60	.58	.56	.54	.52	.51
.35	.78	.74	.70	.66	.64	.61	.59	.57	.55	.53	.51
.40	.82	.78	.73	.69	.66	.63	.61	.58	.56	.53	.52
.45	.85	.81	.75	.71	.68	.65	.62	.59	.56	.53	.52
.50	.88	.84	.78	.74	.70	.67	.63	.60	.57	.54	.52
.55	.91	.87	.81	.76	.72	.69	.65	.61	.58	.54	.52
.60	.94	.90	.84	.79	.75	.70	.66	.62	.59	.54	.52
.65	.96	.92	.87	.82	.77	.73	.68	.64	.59	.55	.52
.70	.98	.95	.90	.85	.80	.75	.70	.65	.60	.55	.53
.75	.99	.97	.92	.87	.82	.77	.72	.66	.61	.55	.53
.80	1.00	.99	.95	.90	.85	.80	.73	.67	.61	.55	.53
.85	1.00	.99	.97	.94	.88	.82	.76	.69	.62	.55	.53
.90	1.00	1.00	.99	.97	.92	.86	.78	.70	.62	.56	.53
.95	1.00	1.00	1.00	.99	.96	.90	.81	.71	.63	.56	.53
1.00	1.00	1.00	1.00	1.00	1.00	1.00	.83	.71	.63	.56	.53

Proportion of Employees Considered Satisfactory = .60

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60	.60
.05	.64	.63	.63	.62	.62	.62	.61	.61	.61	.60	.60
.10	.68	.67	.65	.64	.64	.63	.63	.62	.61	.61	.60
.15	.71	.70	.68	.67	.66	.65	.64	.63	.62	.61	.61
.20	.75	.73	.71	.69	.67	.66	.65	.64	.63	.62	.61
.25	.78	.76	.73	.71	.69	.68	.66	.65	.63	.62	.61
.30	.82	.79	.76	.73	.71	.69	.68	.66	.64	.62	.61
.35	.85	.82	.78	.75	.73	.71	.69	.67	.65	.63	.62
.40	.88	.85	.81	.78	.75	.73	.70	.68	.66	.63	.62
.45	.90	.87	.83	.80	.77	.74	.72	.69	.66	.64	.62
.50	.93	.90	.86	.82	.79	.76	.73	.70	.67	.64	.62
.55	.95	.92	.88	.84	.81	.78	.75	.71	.68	.64	.62
.60	.96	.94	.90	.87	.83	.80	.76	.73	.69	.65	.63
.65	.98	.96	.92	.89	.85	.82	.78	.74	.70	.65	.63
.70	.99	.97	.94	.91	.87	.84	.80	.75	.71	.66	.63
.75	.99	.99	.96	.93	.90	.86	.81	.77	.71	.66	.63
.80	1.00	.99	.98	.95	.92	.88	.83	.78	.72	.66	.63
.85	1.00	1.00	.99	.97	.95	.91	.86	.80	.73	.66	.63
.90	1.00	1.00	1.00	.99	.97	.94	.88	.82	.74	.67	.63
.95	1.00	1.00	1.00	1.00	.99	.97	.92	.84	.75	.67	.63
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.86	.75	.67	.63

Proportion of Employees Considered Satisfactory = .70

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70	.70
.05	.73	.73	.72	.72	.72	.71	.71	.71	.71	.70	.70
.10	.77	.76	.75	.74	.73	.73	.72	.72	.71	.71	.70
.15	.80	.79	.77	.76	.75	.74	.73	.73	.72	.71	.71
.20	.83	.81	.79	.78	.77	.76	.75	.74	.73	.71	.71
.25	.86	.84	.81	.80	.78	.77	.76	.75	.73	.72	.71
.30	.88	.86	.84	.82	.80	.78	.77	.75	.74	.72	.71
.35	.91	.89	.86	.83	.82	.80	.78	.76	.75	.73	.71
.40	.93	.91	.88	.85	.83	.81	.79	.77	.75	.73	.72
.45	.94	.93	.90	.87	.85	.83	.81	.78	.76	.73	.72
.50	.96	.94	.91	.89	.87	.84	.82	.80	.77	.74	.72
.55	.97	.96	.93	.91	.88	.86	.83	.81	.78	.74	.72
.60	.98	.97	.95	.92	.90	.87	.85	.82	.79	.75	.73
.65	.99	.98	.96	.94	.92	.89	.86	.83	.80	.75	.73
.70	1.00	.99	.97	.96	.93	.91	.88	.84	.80	.76	.73
.75	1.00	1.00	.98	.97	.95	.92	.89	.86	.81	.76	.73
.80	1.00	1.00	.99	.98	.97	.94	.91	.87	.82	.77	.73
.85	1.00	1.00	1.00	.99	.98	.96	.93	.89	.84	.77	.74
.90	1 00	1.00	1.00	1.00	.99	.98	.95	.91	.85	.78	.74
.95	1.00	1.00	1.00	1.00	1.00	.99	.98	.94	.86	.78	.74
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.88	.78	.74

Proportion of Employees Considered Satisfactory = .80

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
.00	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80	.80
.05	.83	.82	.82	.82	.81	.81	.81	.81	.81	.80	.80
.10	.85	.85	.84	.83	.83	.82	.82	.81	.81	.81	.80
.15	.88	.87	.86	.85	.84	.83	.83	.82	.82	.81	.81
.20	.90	.89	.87	.86	.85	.84	.84	.83	.82	.81	.81
.25	.92	.91	.89	.88	.87	.86	.85	.84	.83	.82	.81
.30	.94	.92	.90	.89	.88	.87	.86	.84	.83	.82	.81
.35	95	.94	.92	.90	.89	.89	.87	.85	.84	.82	.81
.40	.96	.95	.93	.92	.90	.89	.88	.86	.85	.83	.82
.45	.97	.96	.95	.93	.92	.90	.89	.87	.85	.83	.82
.50	.98	.97	.96	.94	.93	.91	.90	.88	.86	.84	.82
.55	.99	.98	.97	.95	.94	.92	.91	.89	.87	.84	.82
.60	.99	.99	.98	.96	.95	.94	.92	.90	.87	.84	.83
.65	1.00	.99	.98	.97	.96	.95	.93	.91	.88	.85	.83
.70	1.00	1.00	.99	.98	.97	.96	.94	.92	.89	.85	.83
.75	1.00	1.00	1.00	.99	.98	.97	.95	.93	.90	.86	.83
.80	1.00	1.00	1.00	1.00	.99	.98	.96	.94	.91	.87	.84
.85	1 00	1.00	1.00	1.00	1.00	.99	.98	.96	.92	.87	.84
.90	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94	.88	.84
.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.96	.89	.84
1.00	1 00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.89	.84

Proportion of Employees Considered Satisfactory = .90

Selection Ratio

r	.05	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
00	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90	.90
.05	.92	.91	.91	.91	.91	.91	.91	.90	.90	.90	.90
.10	.93	.93	.92	.92	.92	.91	.91	.91	.91	.90	.90
.15	.95	.94	.93	.93	.92	.92	.92	.91	.91	91	.90
20	.96	.95	.94	.94	.93	.93	.92	.92	.91	91	.90
25	.97	.96	.95	.95	.94	.93	.93	.92	.92	.91	.91
.30	.98	.97	.96	.95	.95	.94	.94	.93	.92	.91	.91
.35	.98	.98	.97	.96	.95	.95	.94	.93	.93	.92	.91
.40	.99	.98	.98	.97	.96	.95	.95	.94	.93	.92	.91
.45	.99	.99	.98	.98	.97	.96	.95	.94	.93	92	.91
.50	1.00	.99	.99	.98	.97	.97	.96	.95	.94	.92	.92
.55	1.00	1.00	.99	.99	.98	.97	.97	.96	.94	.93	.92
.60	1.00	1.00	.99	.99	.99	.98	.97	.96	.95	.93	.92
.65	1.00	1.00	1.00	.99	.99	.98	.98	.97	.96	.94	.92
.70	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.96	.94	.93
.75	1.00	1.00	1.00	1.00	1.00	.99	.99	.98	.97	.95	.93
.80	1.00	1.00	1.00	1.00	1.00	1.00	.99	.99	.97	.95	.93
.85	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.98	.96	.94
.90	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	.99	.97	.94
.95	1.00	1.00	1.00	1 00	1.00	1.00	1.00	1.00	1 00	.98	.94
1 00	1.00	1.00	1.00	1 00	1.00	1.00	1.00	1.00	1.00	1.00	.95

APPENDIX C

PUBLISHERS OF REPRESENTATIVE PERSONNEL TESTS

The tests in this Appendix include only a very small number of the personnel tests commercially available. They have been limited almost entirely to tests that have been discussed in this book. Very good additional sources of information about the availability and uses of personnel tests are:

O. K. Buros, ed., *The Fourth Mental Measurements Yearbook* (New Brunswick, N.J., Rutgers University Press, 1953).

This book, in addition to giving publishers and prices, contains critical reviews of each test by one or more competent authorities.

R. M. Dorcus and M. H. Jones, *Handbook of Employee Selection* (New York, McGraw-Hill Book Co., Inc., 1950).

This book gives abstracts of 426 articles dealing with personnel tests that have been published in professional journals. It contains indexes by job, test, and author. The abstracts are factual and non-critical.

In the lists given below, the publisher of each test is coded in the column indicated.

Title	*Code to Publisher (see pp. 537–538)*
Intelligence Tests	
Adaptability Test	12
Army Alpha, Forms A and B, Bregman's Revision	10
Army Alpha, Forms 5 and 7, Well's Revision	10
Army Alpha, Short Form 6	10
Army General Classification Test, First Civilian Edition	12
Benge Employment Tests	9

California Test of Mental Maturity	3
Cattell Culture-free Test	10
Chicago Non-Verbal Examination	10
D.A.T. Abstract Reasoning Test	10
D.A.T. Numerical Reasoning Test	10
D.A.T. Verbal Reasoning Test	10
Ohio State University Psychological Test, Form 21	12
Oral Directions Test: An Intelligence Test on Phonograph Records	10
O'Rourke General Classification Test	11
Otis Self-Administering Tests of Mental Ability	16
Psychological Examination of the American Council on Education	4
Purdue Non-Language Personnel Test	12
Revised Beta Examination (non-verbal)	10
S.R.A. Non-Verbal Form	12
S.R.A. Verbal Form	12
Thurstone Test of Mental Alertness	12
Wonderlic Personnel Test	10

Clerical Tests

E.R.C. Stenographic Aptitude Test	12
D.A.T. Clerical Speed and Accuracy	10
General Clerical Test	10
Minnesota Clerical Test (formerly Minnesota Vocational Test for Clerical Workers)	10
Purdue Clerical Adaptability Test Form A (revised)	17
S.R.A. Clerical Aptitudes	12

Industrial Vision Tests

Keystone Visual Safety Tests	7
Ortho-Rater	2
Sight-Screener	1

Stenographic and Typing Skills

Bennett Stenographic Aptitude Test	10
Bennett Stenographic Proficiency Test	10
Seashore-Bennett Stenographic Proficiency Test (on phonograph records)	10
S.R.A. Typing Skills	12
S.R.A. Dictation Skills	12
Thurstone Examination in Typing	16

Dexterity and Manipulative Tests

MacQuarrie Test for Mechanical Ability	3
Minnesota Rate of Manipulation Test	15
O'Connor Finger Dexterity Test	8
O'Connor Tweezer Dexterity Test	8
Purdue Hand Precision Test	8
Purdue Pegboard	12

Mechanical Tests

Bennett Mechanical Comprehension Test	10
Industrial Training Classification Test	12
Minnesota Mechanical Assembly Test	15
Minnesota Paper Form Board, revised	10
Minnesota Spatial Relations Test	10
O'Connor-Wiggly Block Test	5
O'Rourke Mechanical Aptitude Test—Jr. Grade	10
Purdue Mechanical Adaptability Test	17
Purdue Mechanical Performance Test	8
S.R.A. Mechanical Aptitudes	12
Stenquist Mechanical Aptitude Test	10

Personality Tests

A-S Reaction Study, Beckman's Revision for Business Use	10
Gordon Personal Profile	16
Gordon Personal Inventory	16
Bell Adjustment Inventory	10
Bernreuter Personality Inventory	14
Guilford Inventory of Factors GAMIN	13
Guilford Inventory of Factors STDCR	13
Guilford-Martin Personnel Inventory	13
Guilford-Zimmerman Temperament Survey	13
Humm-Wadsworth Temperament Scale	6
Minnesota Multiphasic Personality Inventory	10
Thurstone Temperament Schedule	12

Interest Tests

Cardall Primary Business Interests Tests	12
Cleeton Vocational Interest Inventory	10
Kuder Preference Record	12
Strong Vocational Interest Blank for Men, Form M	10
Strong Vocational Interest Blank for Women, Form W	10

Trade Tests

Purdue Industrial Mathematics Test	17
Purdue Interview Aids	
Can You Read a Working Drawing?	12
Can You Read a Micrometer?	12
Can You Read a Scale?	12
Purdue Vocational Tests	
Technical Information in Machine Shop	12
Technical Information in Electricity	12
Blueprint Reading Test	12
Purdue Trade Information Test in Welding	17
Purdue Trade Information Test in Carpentry	17
Purdue Trade Information Test in Engine Lathe Operation	17

General

Purdue Reading Test for Industrial Supervisors	17
Purdue Industrial Supervisors Word-Meaning Test	17

PUBLISHERS

1. American Optical Company	Southbridge, Massachusetts
2. Bausch & Lomb Optical Company	Rochester, New York
3. California Test Bureau	5916 Hollywood Boulevard Los Angeles 28, California
4. Educational Testing Service	2½ Chambers Street Princeton, New Jersey
5. Human Engineering Laboratory, Inc.	347 Beacon Street Boston 16, Massachusetts
6. Humm Personnel Service	P.O. Box 1433 Del Valle Station Los Angeles 15, California
7. Keystone View Company	Meadville, Pennsylvania
8. Lafayette Instrument Company	52 By-Pass Lafayette, Indiana

9. Management Service Company	3136 North 24th Street Philadelphia, Pennsylvania
10. Psychological Corporation	304 East 45th Street New York 17, New York
11. Psychological Institute	P.O. Box 1117 Lake Alfred, Florida
12. Science Research Associates	57 West Grand Avenue Chicago 10, Illinois
13. Sheridan Supply Company	P.O. Box 837 Beverly Hills, California
14. Stanford University Press	Stanford, California
15. C. H. Stoelting Company	424 North Homan Avenue Chicago 20, Illinois
16. World Book Company	313 Park Hill Avenue Yonkers 5, New York
17. University Book Store	360 State Street West Lafayette, Indiana

INDEX

A

Abel, T. M., 122, 122n
Ability, 110, 374
  definition, 81
  latent, 79, 81
  mental, accident-proneness, 423
  physical, 434–436
Absenteeism, 9, 34, 284–285, 329, 332
  communications, 355, 356 (*Table*)
  increase of work-time, 456
  labor-management participation, 366 (*Table*)
  morale, 304
Accuracy, 32
Accident-prone employee, 401
  identifying, 425–426
  (*see also* Accident proneness and Accidents)
Accident proneness, 32, 401–420, 427–432
  chance factors, 402
  external factor related to, 409–412
  factors within individual, 401, 402, 406–407
  hospital visits and causative factors, 404–409
  identifying factors related to, 426–427
  job hazards in relations to, 406–409
  muscular speed, 424–425
  perceptual speed, 424–425
  personal variables, 412
  primary causes, 420 (*Fig.*)
  placement of employees, 427–428
  psychological factors related to, 412
  reduction of, 427–432
  steel mill, 404 (*Table*), 405 (*Fig.*)
  vision, 413–419
  working conditions in relation to, 401, 402, 406–407, 412
  work situation, 412
  (*see also* Accidents)
Accidents, 34, 284, 398–432
  age factor, 415–419
Accidents (*cont.*)
  causes, 399–409
  emotional factors, 419–422
  first aid, 401
  first-aid cases, 399
  home cases, 399
  job experience factor, 415–419
  location of medical facilities, 432
  lost time, 399, 401
  medical attention, 432
  mental ability, 422–423
  music, 455
  new employees, 428–429
  personal protection, 429–431
  preventability, 398
  proneness (*see* Accident proneness)
  safety consciousness, 428
  safety devices, 427–432
  steel mill, 399–401, 400 (*Table*)
  types of injury, 431 (*Table*)
  visual skill tests in relation to, 413 (*Table*)
Achievement tests, 79, 86, 185–199, 202–203
  job sampling, 190–195
  machinist, 80
  oral, 186–189
  purpose, 80, 185–186
  uses of, 199–203
  validity of, 185–186
  written, 195–199
Acme Steel Company, merit rating program, 247
Accounting, 174
Action-taking process, 469–473
Adams, S., 450, 450n
The Adaptability Test, 79, 113–114, 113n, 117, 118–120, 121, 126
Adjustor, 174
Adjutant General's Office, 219–221
Adult placement test, 120
A.F.P.T.R.C. (*see* Air Force Personnel and Training Research Center)

Aggression, 338
Aircraft recognition training, 274
Air flow, 449
Air Force Personnel and Training Research Center, 123n
Air Line Mechanics Associations, 500
Airplane-engine-and-propeller plant, 132
Alessi, Salvatore L., 502n, 508, 508n
Allen, E. P., 144, 144n
Allen, L. A., 296, 296n
Alternate-form method, 93
Allport, Gordon, 3, 3n
American Council on Education Psychological Examination, 133
American Engineering Council, 411
American Management Association, 361n, 387n
American Medical Association, 149
  section of industrial health, 148n
  visual efficiency scale, 148–149, 149 (*Fig.*)
American Optical Company, 557
American Psychological Association, 36n, 91n, 133n
American Society of Heating and Air-Conditioning Engineers, 449n
Anderson, R. G., 120, 120n
Anderson, V. V., 422, 422n
Angles, A., 455, 455n
Applicants, 39, 77, 108, 109, 186, 198
  achievement tests, 199
  application form, 66–70
  behavior sampling, 59
  dexterity, 161
  evaluating proficiency of, 65
  factors for job failure, 162
  faking answers, 170, 173, 174–176, 179
  incentives, 324
  individual differences, 15–20
  interviewing, 2
  mental ability, 161
  personal characteristics, 161
    judgment of, 60–61
  personnel tests, 73–74
  references, 70–73
  validating tests, 85
  work history, 62–65
Application forms, 2, 39, 48, 51, 53, 66–70
  development, use of statistical comparison, 67–70
Application forms (*cont.*)
  pertinence, 66–67
  use of template, 70
  value of length 67
Apprentices, 144, 281–283
  measurement of accuracy of, 383
  selection of, 129
  training, 288, 293–294
Apprentice-Training Service, 294n
Aptitude, 86
  measurements of, 110–160
Aptitude index, 68–69
Aptitude tests, 79–80, 81, 110–160, 294
  classes, 110
  clerical, 156
  hosiery looping, 79–80
  mechanical ability:
    group tests, 124–127
    individual tests, 127–129
    nature of, 123–124
    tests in use, 129–134
  mental ability, 110
    nature of, 111–112
    tests, 112–122
  perception, 156–157
  psychomotor ability:
    nature of, 134–136
    tests, 136–145
  purpose, 79–80
  special, 156–157
  testing, 86
  validity of, 157–160
  visual skills:
    nature of, 146–148
    tests, 148–156
Arbitration, 4
Arithmetic mean, 519–521, 524
Armed forces, 274
Armstrong, T. O., 207, 207n
Army Alpha Mental Test, 423
Army officers, 219–221
  pooled ratings, 244–245
Army Ordnance School, 201
Arnold, J. N., 197n
Arps, G. F., 440, 440n
Asbestos workers, 186, 187 (*Fig.*)
Asher, E. J., 113n, 138n
Assorting, 375
Astrology, 2
Atlantic Refining Company, 228n
Atmospheric conditions of work, comfort zones, 447–449

Attitudes, 327–334, 335, 340
  absenteeism, 329–332
  accident factor, 419
  behavior, 332–334
  definition, 302, 303
  job performance, 328–329
  measurement of, 302–326
  merit rating, 330–331
  tardiness, 329–332
  tenure, 329–332
  (*see also* Morale)
Attitude scale, 305–313
  administration, 307–308
  calculation of, 307–308
  construction of, 305–307
Attitudinal variables, 329
Autistic schizoid component, 166
Automotive mechanics tests, 201–202
Aviation Psychology Program, 469n

B

Bachus, J. A., 215n
Baetjer, Anna M., 448, 448n
Baker, job description, 43–44 (*Table*)
Bank employees, 116
Barnes, R. M., 472, 472n, 473, 475n, 476, 476n, 477n
Barre, M. F., 139, 139n
Barthol, R. P., 181, 181n, 182, 183
Bartley, S. H., 436, 436n
Batson, W. H., 264n
Baumgartel H., 331, 331n. 355, 355n, 369, 369n
Bausch & Lomb Optical Company, 557
Bausch & Lomb visual classification and placement tests for industry, 150
Bayroff, A. G., 245, 245n
Beckett, W., 454, 454n
Bedaux units, 276
Bedford, T., 448, 448n
Behavior, 327–334
  attitudes, 332–334
  causation in, 334–335
    of consumer, 7
  frustration, 338–339
  habits, 335
  incentives, 335–338
  motivation, 336–338
  needs, 335–338
  understanding of, 334–339
Benge, E. J., 493, 493n, 494
Bennett-Fry test, 124n
Bennett, George K., 124n, 195, 195n
Bennett Test of Mechanical Comprehension, 79, 124–125, 124n, 125 (*Fig.*), 132, 133
Beranek, L. L., 451
Berdie, R. F., 325
Bergen, H. B., 305, 305n
Bergen scale, 305 (*Table*)
Berkshire J. R., 220, 220n, 221 221n
Berman, I. R., 139n, 142 (*Table*)
Bernberg, R. E., 328, 328n
Bernreuter Personality Inventory, 163–164
Bernreuter, R. G., 164, 164n, 165, 165n
Bernreuter scale, 117
Besnard, G. G., 394, 394n
Bills, A. G., 435, 435n, 437, 437n
Bills, M. A., 12, 12n, 116, 116n, 122, 122n
Bingham, W. V., 36, 36n, 58, 58n, 241n
Birren, F., 444, 444n
Bishop, M. K., 23, 23n
Bittner, R. H., 36, 36n, 245, 245n
Blai, Boris, Jr., 9, 9n
Blackstone Stenographic Proficiency Tests, 194, 194n
Blankenship, A. B., 276, 276n
Blind-positioning movements, 469–471, 470 (*Figs.*)
Blueprint reading test, 197
Blum, M. L., 141, 141n, 323, 323n, 325, 402, 402n
Bookkeepers, 118
Book, W. F., 265, 265n
Boredom, 12, 433, 436, 439, 441
Bottle selectors, placement of, 388
Box manufacturer, 456
Boyd, J. D., 292, 293
Brayfield, A. H., 319, 319n, 327, 327n, 333, 333n
Breakage, 283
Breaks (*see* Rest pauses)
Bregman, E. O., 271n
Brewer, J. M., 162, 162n
Bricklaying, 474–475, 477
Bright, C. J., 169, 169n
Brightness contrast, 442–443, 443 (*Fig.*)
Brogden, H. E., 38, 38n
Brouha, L., 434, 434n, 441, 441n. 478, 478n, 479
Brown, C. W., 151, 151n, 157, 157n
Brown, Robert L., 129n

Bryan, W. L., 477n
Burk, S. L. H., 493, 493n, 494
Burlington Mills Corporation, 144 (*Fig.*)
Buros, Oscar K., 116n, 554
Burton, A., 169, 169n
Burus, R. K., 221, 221n
Buxton, C. E., 134, 134n

C

Caldwell, C. W., 197n
California Test Bureau, 557
Carnegie, Dale, 344
Case, H. W., 129, 129n, 130n
Cashiers, 117
Central tendency measures, 519–524
Chambers, E. G., 423, 423n, 426, 426n
Chain reaction, 301
Change Making Test, 117
Chant, S. M. F., 325
Chapanis, A., 462n
Character and Personality, 424n
Chave, E. J., 305, 305n
Check list rating systems, 216–221
Chesler, D. J., 502, 502n, 511, 511n
Children's tests, 113, 115
Christy, F. C., 175n
Chute, E., 436, 436n
Civil Service classification system, 495
Clarke, W. V., 117, 117n
Classification method, 494–495
Cleeton, G. E., 3n
Cleeton, G. U., 76, 76n
Clerical tests, 116, 156, 194–195
  list of, 555
Clerks, 182
Cleveland Railway Company, 419, 420
Clinical counseling, 183–184
Coch, L., 362, 363n
Cochran, R., 198n
Coe, H. G., 388, 388n
Colby, A. N., 359, 359n
Collective bargaining, 485–486
Collector, 174
College graduates, 172
Command and General Staff College, 245
Committee on Work in Industry of the National Research Council, 440n
Communications:
  ability, 355
  absenteeism, 355, 356 (*Table*)
Communications (*cont.*)
  definition, 347
  effectiveness, 352
  flow, 347
  formal, 347–348
  growth of industry, 346–347
  human relations, 346–360
  informal, 347–348
  methods, 347–353
  oral, 350, 352
  reading-ease of house organs, 359
  strength and weaknesses of media, 348–349 (*Table*)
  value, 353–355, 354 (*Table*)
  viewpoints of management, 350–351 (*Table*)
  written, 350, 352
    readability, 357–360
Communicative ability, 355–356
Community effect on wages, 484, 485
Company policy, 286, 292
Comparable scores, 527–532
Completion form test, 297
Concurrent validity, 92, 94
Conferences, 295, 296, 396
  safety, 417, 428–429
Confidence level, 48
Constant error, 228
Content validity, 92
Cook, D. W., 139, 139n, 387, 387n
Correlation, 534–547
Costs, 4, 5, 25, 48
  equipment, 5
  labor, 30
  living, 485
  material, 5
  overhead, 25, 30
  production, 1, 5, 14, 486
  training, 14, 33
  turnover, 3–4
Cox, K. J., 171, 171n
Cox Mechanical Aptitude Test, 133
Criterion, types of, 33–35
Criterion contamination, 36–37
Criticism, 268
Crockett, W. H., 327, 327n, 333, 333n
Culpin, M., 476n
Cureton, E. E., 171n
Cycloid, 165

D

Dahle, T. L., 352, 352n, 353
Dale, E., 361, 361n, 365, 365n

Daniels, H. W., 490, 490n
Data, statistical, graphic representation of, 514–519
D.A.T. Mechanical Aptitude Tests, 124n
Davis, H. A., 207, 207n
Davis, L. E., 471, 471n, 472
Davis, M. K., 506, 506n
Daydreaming, 12
Decibels, 450, 453
  noise sources, 451 (*Fig.*)
Decision-making process 465–469
Decisions, 467–468
Demonstrations, 395
Depression, 166
  accident factor, 419
Depression, 1930–1935, 77–78
Depression scale, 169
Depressive cycloid component, 165
Department of the Army, 252n
Department store, women salesclerks, 49–50
  job performance in relation to job satisfaction, 328
Dependability, 27
Depth perception, 146–147
Descriptive statistics, 513
Detter, Howard M., 123, 123n
Dexterity:
  tests, 93, 140–144
    list of, 556
  trade apprentices, 144–145
  (*see also* Finger dexterity *and* Manual dexterity)
Dexterity and muscular co-ordination tests, 142. (*Table*)
Diagnostic Interviewer's Guide, 56, 57 (*Fig.*), 62
Dickson, W. J., 368, 368n
Differential production, 76
Disabled:
  job analysis, 46 (*Fig.*)
  jobs for, 47–48
Disciplinary action, 9
Discrimination, 487
Discussions, 295
Dorcus, R. J., 554
Dotting test, 140 (*Fig.*)
Drake, C. A., 316, 316n, 424, 424n, 425
Drake, L. E., 169, 169n
Driver, R. S., 222n, 228, 228n, 248
Drucker, P. F., 369, 369n
Dudek, F. J., 134, 134n
Dunlap, J. W., 53n

E

Educational Testing Service, 557
Efferson, Carlos A., 296
Electrical-fixture assemblers, 24–25
Electrical inspectors, placement of, 287–288
Electricians' test, 197
Emotions, 419–422
Employability, 66–74
Employee attitudes (*see* Attitudes, Measurements, *and* Opinions)
Employee comparison systems, 213–216
Employee opinion surveys (*see* Opinion surveys)
Employee participation, 367
  areas of, 363 (*Table*)
  desirability of, 360–361
  human relations, 360–368
  types of, 361–367
Employees:
  accidents, 428–429
  evaluation of (*see* Merit rating)
  merit rating interview, 206–208
  preference for music, 454–455
  production efficiency, 455–457
  revealing rating information to, 234–235
  training, 2–3, 251–301
  weighting traits, 237–242
Employee self-improvement, 206–208
Employment managers, 39, 81–82, 84–85, 88, 101, 108, 161, 162
  achievement test, 198–199
  sources of information for, 116
  (*see also* Personnel managers)
Employment methods:
  applicants, information on, 52–74
  application form, 39, 66–70
  early, 1–3
  flow chart, 40 (*Fig.*)
  improvements, 76
  individual differences, 31–32
  interview, 39–40, 52–74
  office procedures, 39–41, 75–76
  personnel specifications, 41–52
  references, 39, 71–73
  selection procedure, 40 (*Fig.*)
  tests, 39, 40–41, 73–74, 76, 81, 120
Employment offices, 39–41, 53, 75–76

Employment stability, 328, 329 (*see also* Tenure)
Employment tests, 533–534 (*see also* Personnel tests)
Engine lathe operation test, 198
England, A. O., 359, 359n
Environment, 31
Epileptoid component, 165
Equipment design, 460
Essex County Vocational School, 423
Evans, C. E., 313n, 315
Evans, R. N., 383, 383n, 395, 395n
Exit interviews, 9, 316–317
Expectancy chart, 84, 84 (*Fig.*)

F

Factor analysis, 111, 111n
Factor comparison method, 493–494
Fact tests (*see* Flanagan Aptitude Classification Tests)
Fair Labor Standards Acts, 484
Farmer, E., 423, 423n, 426, 426n, 476n
Fatigue, 433, 436, 439, 440, 458 (*see also* Industrial fatigue)
Favorability index, 220
Fear, as cause of accidents, 419
Fear, R. A., 54, 54n, 56
Featherbedding, 369
Federal Committee on Apprenticeship, 294–295
Feeble-minded, adaptability test for, 122
Feinberg, M. R., 253, 253n, 254, 274, 274n, 294, 294n
Ferree, C. E., 150n
Filing, 174
Finger dexterity, 80, 135, 161
  O'Connor test for, 17–18
Firing, 162
  early procedures, 3
  reasons for, 5
  (*see also* Layoffs *and* Termination)
Fitts, P. M., 467, 467n, 469, 469n, 470
Fixation, 338–339
Flanagan, J. C., 164, 164n, 221, 221n
Flanagan Aptitude Classification Tests, 125–126
Fleishman, E. A., 300, 300n
Flesch, R., 356, 356n, 357
Flesch Reading Ease measurement, 356–358
Flinn, R. H., 400n
Floor, L. G., 345, 345n
Flow chart, 40 (*Fig.*)
Forbes, J. W., 468, 468n
Forced-choice check list, 219–221
Forced-choice technique, 170
Forced-choice test, fallability of, 176–178
Forced distribution system, 215–216
Foremen, 4, 303, 365
  aid to jobs analysis, 490–491
  changes in function, 295
  disputes, 4
  employee attitudes, 339
  interest test, 174
  merit-rating interview, 206–208
  morale, 304
  training by, 251
  training of, 4–5
Fork-lift operators, 193–194
  standard driving course, 193 (*Fig.*)
Fraser, A., 449, 449n
French, J. R. P., Jr., 363, 363n
Frequency distribution, 514–518
Friedman, Gabriel, 123, 123n, 124, 124n, 134, 134n, 136, 136n
Frisby, C. B., 270, 270n
Frustration, 338–339
Fryer, D. H., 253, 253n, 254, 272, 272n, 273, 274, 274n, 294, 294n
Funk, F. E., 347n, 350n, 351

G

Garner, W. R., 462n
Gates, D. S., 416, 416n
General education, 202–203
General Motors Corporation, 8
  employee research section, 313
  "My Job and Why I Like It" contest, 311, 313–315
  theme rank order, 311 (*Table*)
Ghiselli, E. E., 157, 157n, 181, 181n, 182, 183
Giese, W. J., 151n, 175n, 331, 331n, 332, 332n
Gilbreth, Frank B., 474, 474n, 475n, 476n, 477, 477n, 478, 478n, 479
Gilbreth, Lillian M., 475n, 476n
Globe Optical Company, 151n
Glorig, A., 452, 452n, 453
Goals (*see* Incentives)
Goppert, H. R., 197n
Gordon, L. V., 177, 177n, 178
Gordon Personal Profile and Personal Inventory, 170, 176–177

Gorsuch, John H., 9, 9n
Graphic rating scale, 210
Graphologist, 2
Graphology, 2
Great Britain, 475n
Green, H. J., 139n, 142 (*Table*)
Greenly, R. J., 13n, 17n, 122, 122n, 141, 141n, 190, 190n, 210n, 211n, 280, 280n, 299
Greenwood, M., 403, 403n
Grievances, 4, 9
  airing of, 316–317
Griffin, W. B., 198n
Griffith, J. W., 458, 458n
Griffitts, C. W., 136, 136n
Group:
  acceptance, 326
  behavior, differences in, 368–370
  differences, 327
  feeling, 326
  goals, 369–370
  participation, 363 (*see also* Employee participation)
  tests, 79, 112
  testing, 138, 138n, 139
  unity:
    costs of, 371
    development of, 370–372
    differences in group behavior, 369–370
    human relations, 368–372
    job performance, 371–372, 372 (*Table*)
    supervisors, 371
    turnover, 371
Gruber, Alin, 114n
Guided interview, 315–316
Guilbert, H., 430n
Guilford, J. P., 22n, 87n, 103, 166, 166n
Guilford, R. B., 166, 166n
Guilford Series of Personality Tests, 166–168

H

Habbe, S., 347, 348, 348n, 349n, 350n, 353, 354
Habersat, R. W., 48, 48n
Haggard, W. W., 459
Haggerty, Helen R., 245, 245n
Halo effect, 213, 222–228, 248
Hanawalt, N. G., 165, 165n
Hanman, Bert, 48n
Hardt, E. F. E., 317, 317n
Harrell, T. W., 119, 119n, 134, 134n
Harris, F. J., 312, 312n
Hartmann, G. W., 22, 22n, 23, 23n
Harwood Manufacturing Company, group participation study, 362–364
Hathaway, S. R., 166n, 168n
Hawthorne Plant (Western Electric Company), 135n
Hay, Edward N., 117, 117n, 118, 493, 493n, 494
Hayes, Elinor G., 143, 143n
Hayes Pegboard, 105, 135, 136, 137 (*Fig.*), 138, 143
Heinrich, H. W., 398, 398n
Henig, M. S., 423, 423n
Heredity, physical traits determined by, 31
Herrington, L. P., 446, 446n
Hersey, R. B., 321, 321n, 322, 419, 419n, 420, 420n, 459
Hewes, A., 416, 416n
Hick, W. E., 467, 467n
Highland, R. W., 220, 220n, 221, 221n
Hill, Lee H., 485, 485n, 490n
Hill, W. B., 197n
Hinckley, E. D., 307n
Hines, Mildred, 17n, 142 (*Table*)
Hiring, 82, 162
  early methods, 2–3
  (*see also* Employment methods)
Hobbies, 171
Hobbs, G. E., 400, 400n
Hoffman, A. C., 444, 444n
Holmes, F. J., 119, 119n
Hook, C. R., 485, 485n, 490n
Hoppock, 317, 317n, 318, 319n, 322, 322n, 323
Hosiery loopers:
  consistency of production, 21–22, 21 (*Fig.*), 29 (*Table*)
  individual differences:
    effect of training on, 21–22
    experience and, 22
    productivity, 25–26
  productivity distribution, 26 (*Fig.*)
  selection of, 31–32
  shifts in productivity, 28
Hosiery looping, 79, 86
  job experience and production, 278 (*Fig.*)
  learning curve, 21 (*Fig.*), 262

Hosiery looping (*cont.*)
  learning plateau, 21 (*Fig.*), 291
  training, 291
  training aid, 292 (*Fig.*)
Hosiery mill menders, 105
Hours of work, 455–457
House organs, 359
Houser, J. D., 321, 321n, 322
Hovland, C. I., 56
Howard, A. H., 509, 509n
Howard, R. R., 419, 419n
Hull, Clark, L., 98n
Human behavior, 6, 398 (*see also* Behavior)
Human body, and motion economy, 476–481
Human engineering, 6–7, 460–481
  action-taking process, 469–473
  blind-positioning movements, 469–471
  decision-making process, 465–469
  determining size of handwheels, 471
  equipment, 460
  information-receiving process, 462–465
  instrument design, 463–465
  methods analysis, 473–481
  motion and time study, 473–481
  motion economy, 475–481
  non-machine system, 461–462
  purpose of, 460–461
  visual displays, 463–465, 466 (*Fig.*)
  work methods and, 460–481
  work space, 471–473
Human Engineering Laboratory, Inc., 557
Human incentives, 335–338
Human relations, 327, 340–372
  communications, 346–360
  employee participation, 360–368
  group unity, 369–372
  job performance, 341–343
  program, 340
  supervision, 340–346
Hume, C. D., 16n, 196n
Humes, J. F., 455, 455n
Humidity, 445–449
Humm, D. G., 165n
Humm Personnel Service, 557
Humm-Wardsworth Temperament Scale, 165–166
  faking answers, 175 (*Table*)
Hunt, H. C., 162, 162n
"Hunt-and-peck" typing system, 290
Hutcheon, N. B., 133, 133n
Hypochondriasis scale, 169
Hysteroid component, 165–166

I

Illness, 456
Illumination, 441–445
  brightness contrast, 442–443, 443 (*Fig.*)
  optimum levels, 444–445
  visual acuity, 441–442
Impairment, 436
Impulsiveness, 419
Inaba, Kay, 114n
Incentives, 267–268, 337–338, 339, 439–440
  employee attitude, 321–326
  financial, 320–321
  job security, 268
  non-financial, 487
  prediction about, 324
  promotion, 268
  rank order of factors, 325 (*Table*)
  wage plans, 486–487
Individual differences, 11–38, 39, 75–76, 278, 289–290, 336
  bases of, 30–32
  effect of training upon, 20–24
  electrical-fixtures assemblers, productivity of, 24–25
  evaluation, 15–20
  finger dexterity, 17–18
  heredity and environment as causes, 30–32
  hosiery loopers, 25–26
  importance of, 30
  interest areas, 19
  job performances, 24–30
  job satisfaction, 30
  magnitude, 11–23
    increased by training, 22
  merit ratings, 27–28
  productivity, 11–15, 24–27
  significance in industry, 11–38
  stability of job performance, 28–29
  standard employment procedure, 31–32
  technical knowledge, 16–17
Individual mental ability tests, 112
Individual tests, 79
Inductive reasoning, 111

Industrial Accident Prevention, 398n
Industrial employee (*see* Employee)
Industrial engineering, 475
Industrial engineering department, 289
Industrial engineers, 45
  techniques of job analysis, 473–481
Industrial fatigue, 438 (*see also* Fatigue)
Industrial inspection (*see* Inspection)
Industrial management:
  wage payment plan, 267–268
  (*see also* Management)
Industrial mathematics test, 197–198
Industrial merit-rating (*see* Merit-rating)
Industrial psychologist, 32
  function, 75–76
  role, 5
  training employee, 6
Industrial psychology, 1
  contribution, 74
  measurement devices, 6
  personnel problems, 6
  personnel research, 6
  use of, 5, 9–10
Industrial Psychology, 23n, 403n, 458n
Industrial relations trouble, 9 (*see also* Grievances)
Industrial safety (*see* Accident proneness *and* Accidents)
Industrial testing programs, 113
  origin of, 77–78
Industrial Training Classification Test, 126
Industrial training program (*see* Training)
Industrial vision tests (*see* Vision tests)
Industry:
  human relations, 340–372
  social responsibility, 3, 14, 302
Information-receiving process, 462–465
  visual displays, 463–465
Inhibition, 271
Initiative 27
In-service training, 286
Inspection:
  accuracy in use of precision instruments, 380–384, 382 (*Fig.*), 383 (*Fig.*)
  in bottle industry, 388
  in electrical industry, 387–388
  human skills, 373–374
  limit sample, 391, 393
Inspection (*cont.*)
  measuring accuracy of, 374–380
  methods, 388–395, 397
  movement of product, 394
  in tin-plate industry, 374–380
    methods of, 389–391
    training inspectors, 395–397
  work area layout, 394
Inspection program, 397
Inspection test, 157
Inspectors, 442
  accuracy in use of precision instruments, 380–384
  in bottle industry, placement of, 388
  in electrical industry, placement of, 387–388
  measuring accuracy of 374–380
  placement of, 384–388
  selection of, 397
  in tin-plate industry:
    accuracy measurement, 375–380
    placement of, 384–387
    training of, 395–397
  training of, 395–397
  visual skill, 145
Institute for Social Research, 341n, 345n
Instrumental tests, 79
  punch press test, 80
Insurance, 2
  unemployment, 13–14
Intellectual ability, 158
Intelligence, 92, 110, 161
  accident-proneness, 423
  work performance, 12–13
  (*see also* Mental ability)
Intelligence quotient, 112–113, 114
Intelligence tests, list of, 554–555
Interchangeability of tests, 80–81
Interest scale, 169
Interest tests, 79, 171–174
  list of, 556
  purpose, 171
  (*see also* Personality tests)
Interference, 271
International Business Machines Corporation, 46, 47
International Harvester Company, 342
Interstate commerce, 14, 484
Interview, 39–74, 206–207
  application form, 66–70
  deficiencies, sources of, 53
  determining attitudes, 329

Interview (*cont.*)
  determining supervisors' attitudes, 341
  exit, 9, 316–317
  form, 54–56
  guided, 315–316
  methods, 53–74
  opinion surveys, 315–316
  patterned, 54–58
  preliminary, 4
  references, 70–73
  standardized, 54–58
  structured, 54–58
  value of, 53
Interview Aids, 197
Interviewee (*see* Applicant)
Interviewers, 47, 49
  application form, 66–70
  evaluating proficiency of applicant, 65
  evaluating work history, 62–65
  form, 54–56
  job analysis, 42–48
  judgments, 41–42, 59–61
  personnel tests, 73–74
  planning, 54
  procedure, 53–58
  ratings, 61–62
  references, 70–73
  sampling applicants, behavior, 59
  statistical analysis, 48–50
  template, use of, 70
  training of, 2, 58–59
    guide to, 58–59 (*Table*)
  work history, importance of, 62–65
Interviewers' Aid No. 1, 63n
Interviewers' rating scale, 62–63, 63n
Interview procedures, 54
Interview score, 56
Introduction to General Psychology, 113n
Introversion, 92
Inventories (*tests*), 79
  purpose, 81
I. Q. (*see* Intelligence quotient)
Ivens, Frank C., 123, 123n, 124, 124n, 134, 134n, 136, 136n
Ivy, A. C., 456, 456n

J

Jackson, J. J., 269n
Job analysis, 42–48, 389–393, 395
  example of, 43–44 (*Table*)
  preparation of, 45–47
Job analysis (*cont.*)
  for training needs, 256–261
  use of eye camera, 395
  use of motion picture, 389–391, 392 (*Fig.*), 393
  wages, 489
  (*see also* Job evaluation)
Job description check list, 496–497
Job descriptions, 42, 491
Job duties, 42
Jobe, F. W., 151n
Job evaluation, 6, 44, 482–512
  committee, 490, 492, 494
  job descriptions, 491
  program, 488–492
    conversion of job evaluation into wages, 491
    cooperation of supervisors, 490–491
    installation, 489–492
    operations, 489–492
    out-of-line jobs, 492
  statistical studies, 501–512
  systems, 492–497
    classification method, 494–495
    establishing a wage scale, 497–500
    factor comparison method, 493–494
    job description check list, 496–497
    NEMA system, 495–496
    point method, 495–496
    ranking method, 493
  wages, 489
  weighting of items, 500–501, 501 (*Table*)
  (*see also* Job analysis)
Job experience, and production, 277–279
Job-group method, 92
Job groups, relative validity index, 159 (*Fig.*)
Job-group test validation method, 89
Job interest, 162
Job miniature test, 190
Job performance:
  criterion of, 83
  employee attitudes, 328–329
  illumination, 326
  individual differences, 24–30
  merit ratings, 27–28
  music, 455
  noise, 450–452
  prediction of, 81
  rest pauses, 457–459

- Job performance (*cont.*)
  - stability, 28–29
  - training method, 266–267
  - visual acuity, 441–442
  - wage payment plan, 267–268
  - working conditions, 433–459
- Job proficiency, criteria of, 32–38
- Job requirements, 39
- Job sample tests, 375–385
- Job satisfaction, 30, 317–326, 328
  - improved by supervisor, 344
  - individual differences, 11–13
  - measurement of, 317–326
  - value fulfillment, 319–326
- Job satisfaction index, 319 (*Table*)
- Job security, 268, 326
- Job seeking, influence on test answers, 174–175
- Job specification, example of, 43–44 (*Table*)
- Job specialities, 269
- Job success, prediction of, 73
- Job termination (*see* Firing, Layoffs, *and* Termination)
- Johnsgard, J. W., 465, 465n, 466
- Johnson, D. L., 132, 132n
- Johnson, H. M., 105, 105n
- Jones, M. H., 554
- Jordan, B., 54, 54n, 56
- Journeymen, 186, 383
- Judgments, 41–42, 60–61, 92
- Jurgensen, C. E., 139, 139n, 176n, 217, 217n, 308, 308n, 324, 324n
- Jurgensen Classification Inventory, 176

K

- Katz, D., 330, 341, 341n, 345, 345n, 369, 369n
- Keating, E., 62, 62n
- Kelly, E. L., 228, 228n
- Kelly, Martha L., 393n
- Kelly, R. W., 285, 285n
- Kephart, N. C., 214n, 243n, 394, 394n, 414n
- Kerr, W. A., 330, 330n, 332, 332n, 407, 407n, 408, 426, 426n, 454, 454n, 455, 455n, 458, 458n
- *Keystone Telebinocular Test*, 384, 384n
- Keystone View Company, 557
- King, Pearl H. M., 34n
- Kirkpatrick, F. H., 454, 454n
- Kitson, H. D., 268, 268n
- Knauft, E. B., 218
- Knight, F. B., 3n, 113n, 265, 265n
- Knowles, A. S., 207, 207n
- Kohn, H. A., 242, 242n
- König, A., 442, 442n
- Kuder, G. F., 173, 173n
- *Kuder Preference Record*, 19, 19n, 172–173
  - profiles, 19 (*Fig.*)
  - validity for job tenure, 179
- Kuhn, H. S., 148n, 151n
- Kuntz, J. E., 442, 442n
- Kurtz, A. K., 171, 171n

L

- Labor, 162
  - production, factor of, 1
  - shortage and tests, 77–78
  - (*see also* Organized labor *and* Unions)
- Labor disputes, 4
  - wage demands, 482
- Labor grades, 498–499, 499 (*Fig.*)
- Labor-management cooperation (*see* Employee participation)
- Labor turnover, 9, 14
  - exit interview, 316–317
  - morale, 304
  - reasons for, 12–13
  - in relation to work hours, 456
  - training, 285
  - value of testing, 78
  - (*see also* Turnover)
- Lackland Air Force Base, 123n
- Lafayette Instrument Company, 557
- Langdon, J. N., 322, 322n, 325, 455, 455n
- Laseau, L. N., 313n, 315
- Latent ability, 79, 81
- Lauer, Jeanne, 358, 358n
- Lauru, L., 438, 438n, 441n, 478, 478n, 480
- Lauru platform, 438, 441, 478
- Lavos, G., 47–48, 47n
- Lawshe, C. H., 89, 89n, 90, 90n, 91 (*Fig.*), 243n, 257, 258, 342, 342n, 343, 380, 380n, 383, 383n
- Lawshe, C. H., Jr., 13n, 113n, 117, 117n, 118, 118n, 122, 122n, 126n, 131n, 197n, 214n, 281, 281n, 282, 284, 284n, 482, 502, 502n, 503, 503n, 506, 506n, 507n, 508, 508n, 509, 509n
- Laycock, S. R., 133, 133n

Layoffs:
  merit ratings in, 206
  personality factors in, 162
  (see also Firing *and* Termination)
Leadership, 4
Leadership climate. 300, 301
Learning:
  curves, 276, 277 (*Fig.*), 281–283, 282 (*Fig.*)
  interference, 271
  job specialities, 269
  knowledge of results, 265–267
    as influence on job performance, 266 (*Fig.*)
  motivation, 267–269
  practice distribution, 271–272
  principles of, 265
  transfer of training, 269–271
Learning curve, 261–263
Learning plateau, 21 (*Fig.*), 261, 264, 265, 291
Learning time, 33–34
Lee, C. A., 268, 268n
Legislation:
  social, 13–14
  wages, minimum, 14, 34
  (see also Fair Labor Standards Acts)
Leisure-time activities, 171
Lenkart, R. H., 493n, 505n, 506
Letters of recommendation, value of, 71–72, 71 (*Table*)
Level of confidence, 48
Life Insurance Agency Management Association, 69
Life insurance sales managers, 171
Life insurance selling:
  aptitude index, 69
  prior experience and, 64 (*Table*)
Life insurance underwriters, development of application form for, 68
Likert, R., 309, 309n, 310
Likert method, 319
Likert Internationalism Scale, 310, 319
Limit sample, 391, 393
Lindahl, L. G., 258, 258n, 260, 261, 266, 266n, 283n
Link-Belt Company, 57
Lipmann, O., 416, 416n
Locke, N., 241, 241n
Long, W. F., 2, 2n
Longstaff, H. P., 176, 176n
Luckiesh, M., 442, 442n, 443, 444, 444n
Lull, P. E., 347, 350n, 351
Lytle, C. W., 493, 493n, 501, 501n

M

McClelland, D. C. 164, 164n
MacCoby, N., 330, 341, 341n, 345, 345n, 369, 369n
McCollom, I. N., 134, 134n
McComb, H. G., 16n, 196n, 197n
McCord, William S., 296
McCormick, E. J., 188, 188n, 189n, 214n, 215n, 243n, 462n
McDill, J. A. 477n
McGehee, W., 279, 279n, 459, 459n
McGregor, D., 325n
Machine operation, 133
Machine shop test, 196
Machinist, achievement test 80
McKinley, J. C., 166n, 168n
McMurry, R. N., 56, 56n, 57, 58, 116, 116n 132, 132n
McNelly, G. W., 437, 437n
MacQuarrie, T. W., 139n
MacQuarrie Test for Mechanical Ability, 139, 139n
Mahler, W. R., 252, 252n, 275 275n
Maier, N. R. F., 302, 302n, 336n, 338, 338n, 362, 362n, 365, 365n, 367, 367n
Maintenance jobs, 133
Maleski, A. A., 502n, 507, 507n
Management, 4, 83, 178, 372, 384
  accidents 401
  communications, 346–360
  determining attitudes and opinions, 304
  early attitude toward labor, 1
  employee attitudes, 302–326, 333
  employee relations, 346–347
  employee incentives, value of, 325–326
  group feeling, 326
  human relations, 340–372
    motives, 334
  industrial psychologist:
    acceptance of, 7–10
    use of 6
  industrial psychology, hesitancy to use, 9–10
  labor, attitude toward, 14
  merit-rating, administrative uses of, 205–206
  paternalism, 422

Management (*cont.*)
  personnel activities, 7
  placement, 427
  psychological methods, acceptance of, 7 (*Table*)
  psychology, use of, 7 (*Table*)
  rest pauses, 459
  safety programs, 427–432
  selection of supervisors, 343
  social obligation, 3, 14, 302
  training of, 272, 294–301
  training of employees, 251–301
  training program, 287, 397
  treatment of injuries, 432
  use of experts for training, 251
  use of merit-rating, 204
  value system of employees, 319–326
  wages, 483–484
  work hours, 455
Management consultant firm, 304
Management Service Company, 558
Management training, 288, 294–301
Manager, 478n
Manhattan High School for the Women's Garment Trades, 122
Manic cycloid component, 166
Manic-depressive psychosis, 419
Manipulative tests, lists of, 556
Man-machine system, 461–462
  functions of man, 461 (*Fig.*)
Mann, F., 331, 331n, 355, 355n, 369, 369n
Manual dexterity, 117
Mapel, E. B., 286, 286n
Marbe, K., 403, 403n
Marion, Virginia, 362
Maritz, J. S., 402, 402n
Marrow, A., 362, 362n, 363, 365
Mass production, 294
Mayo, Elton, 326, 326n
Mayo, T. B., 458, 458n
Mean, 547
Measurement:
  of attitude, 302–326
  of attitudes and opinions, attitude scales, 305–313
  of job satisfaction, 317–326
  of morale, 302–326
  of opinions, 304–317
Measurement devices, 6
  attitude questionnaires, 6
  personnel tests, 6
  rating procedures, 6
Measurement methods, 261 (*Fig.*), 286–287 296–297
Mechanical ability, 110
  factors, 123–124
  meaning of, 122–123
Mechanical aptitude, 123, 124
Mechanical tests, list of, 556
Median, 521–524
Melrose Hosiery Company, 292, 293
Memory, 111, 296
Memory image, 393
Mengelkoch, R. F., 198n
Mental ability, 110
  accident-proneness and, 423
  nature of, 111
Mental ability tests:
  Flanagan aptitude classification, 115
  individual, 112
  motor, 115
  omnibus, 111
  Otis self-administering, 112–113
  perception, 115
  Purdue non-language personnel, 114
  sources of information, 116
  space, 115
  SRA primary mental abilities, 114–115
  supervisors, selecting, 118–120
  Wonderlic personnel, 113–114
Mental tests, 13
Merit-rating, 6, 204–250
  Acme Steel Company program, 247
  administrative uses, 205–206
  case study, 247–251
  check list rating systems, 216–221
  combining, 245–246
  consistency of, 246–247
  constant error, 228
  critical incident technique, 221–222
  definition, 204
  development, 204–205
  disadvantages, 221–235
  distribution of, 28 (*Fig.*)
  effect of rater's memory, 246–247
  employee attitudes, 330–331
  employee comparison systems, 213–216
  employee self-improvement, 206–208
  factors of, 27
  halo effect 222–228
  improvements, 247–250
  individual differences, 27–28

Merit-rating *(cont.)*
  influencing factors, 228–234
    adjustment of, 235–237
  informing the employee, 234–235
  items, 211 *(Table)*
  layoff, 204–250
  merit compared with seniority, 208–221
  multiple-step rating scale, 212 *(Fig.)*
  pooled ratings, 244–245
  purposes, 205–208
  rating scales 210–213
  reliability, 242–244
  supervisors, 204
  training rater, 248–250
  use of several raters, 244–246
  validity, 242–244
  weighting of traits, 237–241
Merit-rating chart, difficulties of weighting, 238–241
Merit-rating systems, rating scales for, 210–213
Methods analysis, 473–481
Methods improvement, 45
Metropolitan Life Insurance Company, 419
Micrometer, accuracy in use of, 383
Miles, G. H., 268, 268n, 456, 456n
Miles, M. C., 496, 496n
Miller, F. S., 457, 457n
Mind-wandering, 12
Miniature Punch Press Test, 190–193, 191 *(Fig.)*
  use of, 192–193
  validity of, 190–192
  (see also Job sample tests)
Minnesota Clerical Test, 156
Minnesota Mechanical Assembly Test, 127–128, 128 *(Fig.)*
Minnesota Multiphasic Personality Inventory, 168–169
Minnesota Paper Form Board Test, 84 *(Fig.)*, 131, 132, 133
Minnesota Rate of Manipulation Test, 139, 142 *(Table)*, 384–385, 386
Minnesota Spatial Relation Test, 129
Mitchell, J. H., 396n
Mintz, A., 402, 402n, 403, 403n
Mode, 523–524
Monroe, W. H., 252n, 275, 275n
Moore, B. V., 58, 58n
Moore, H., 120, 120n, 207, 207n
Morale, 331, 332
  definition, 303–304
  factors influencing employees', 319–326
  group feeling, 326
  measurement of, 302–326
  (see also Attitudes)
Morgan, C. T., 462n
Morse, N. C., 330, 341, 341n, 345 345n, 369, 369n
Mosby Company, C. V., 149n
Mosel, J. N., 50n, 65, 65n, 66, 72, 72n, 73, 73n
Moss, F. K., 444, 444n
Mossin, A. C., 328, 328n
Motion and time study, 45, 289, 473–481
Motion economy, 473–481
  related to human body, 476–481
Motion pictures:
  in job analysis, 389–391, 392 *(Fig.)*, 393, 394–395
  use in training, 396
Motion study *(see* Motion and time study)
Motivation, 267–269, 336, 340
Motor ability, 158
Motor test, eye-hand coordination, 115
Multiple-choice test, 297
Multiple-step rating scale, 210, 212 *(Fig.)*
Münsterberg, Hugo, 1, 1n
Murray, H. M. L., 456, 456n
Music, 454–455
Musser, W., 239n
Mutual-choice group, 371
Myers, C. S., 456, 456n

N

Nagle, B. F., 38n, 342, 342n
National Conservation Bureau, 47n
National Electrical Manufacturers Association, 495
  job evaluation *(see* NEMA system)
The National Institute of Industrial Psychology, 475n
National Labor Relations Act, 5
National Metal Trades Association, 495
National Society for the Prevention of Blindness, 148n
National War Labor Board, 482
  wage stabilization section, 484

Needs, 335–339
Negative incentive 337
Negative transfer, 271
NEMA system, 495, 496 (*Table*), 500–503, 507–508
Nervousness, 419
New York State Commission on Ventilation, 448 448n
Noise:
  control of, 454
  effect on hearing, 452–453
  effect on job performance, 450–452
  reduction of, 450–452
Normal component, 165
Normal distribution. 518
Northwest Airlines, Inc., 500, 500n
Norvell, L., 265, 265n
Nourishments, 459
Nuckols, R., 331, 331n
Number facility, 111

O

Occupational Research Center, 152
O'Connor, Johnson, 17n, 128n, 142 (*Table*)
O'Connor Finger Dexterity Test, 17–18, 135, 136, 137 (*Fig.*), 138, 141, 142 (*Table*), 143
  frequency distribution of scores, 18 (*Fig.*)
O'Connor Tweezer Dexterity Test, 142 (*Table*)
O'Connor Wiggly Block Test, 128
O'Keefe, R. W., 297, 297n
Oliver, J. A., 509, 509n
Omnibus mental ability tests, 111
Onarheim, J. I., 316, 316n
On-the-job training, 288, 289–293
Opinion questionnaire, 310 (*see also* Opinion survey)
Opinions:
  definition, 303
  measurement of 304–317
  (*see also* Attitudes *and* Measurement)
Opinion surveys, 305, 310–315
  administration of, 312–313
  analysis of employee communications, 314–315
  employee interviews, 315–316
  questionnaire, construction of, 312
  theme analysis, 313–315
Optical manufacturing employees, 48–49
Optometrist, 146
Oral achievement tests, 186–189
  industrial uses, 188–189, 189 (*Fig., Table*)
Oral trade questions, 186–188
  factors affecting validity, 187–188
  validation of, 187 (*Fig.*)
Organized labor, 3, 5, 9, 485 (*see also* Labor *and* Unions)
Orientation training, 287
Ortho-Rater, 150–152, 151 (*Fig.*)
  tests, 152
Osborne, E. E., 411, 411n
Otis, Arthur S., 112, 112n
Otis, J. L., 38n, 130, 130n, 493n, 505, 505n, 506
Otis Self-Administering Tests of Mental Ability, 16, 79, 112–113, 112n, 117, 133, 136
  raw score on, 16 (*Fig.*)
Output restrictions, 368–369
Owen, E. B., 459, 459n
Owen, H. F., 16n, 196n, 197, 197n
Owens-Bennett Test, 124n
Owens, W. A., 133, 133n

P

Paired comparison system, 213–215, 393
Panlasigui, I., 265, 265n
Paper and pencil tests, 79, 112, 139, 163
Paranoia scale, 169
Paranoid schizoid component, 165
Parker, B. T., 48, 48n
Paternalistic system, 325
Paterson, D. G., 62, 62n, 139n, 358, 358n, 359n
Patterned interview, 54
Pay, 321–322, 323, 337 (*see also* Wages)
Peatman, John G., 513n
Pedagogical Seminary, 265n
Pension plans, 313 (*see also* Retirement)
Percentiles, 532–534
Perception, 465
  tests, 115, 155–156
Performance index, character of, 37–38
Performance test, fork-lift operators, 193–194
Personality tests, 87, 163–171
  difficulties of, 162–163

Personality tests (*cont.*)
  faking answers 170, 173, 174–178
  limitations, 174–178
  list of, 556
  purpose, 81
  use of, 178–179, 182–183
  uses for, 183–184
  validation, 179–181
  validity for employment, 181–183
Personality traits, 163
  cause for layoffs, 162
  evaluation of, 2–3
  promotion, 162
  types of, 170
Personnel managers, 13, 39, 79, 172
  exit interview, 316–317
  selection, 192–193
  (*see also* Employment managers)
Personnel office, apprentice training, 294
Personnel policy, 206 (*see also* Human relations)
Personnel program, 6, 9, 319–326
Personnel psychologist, 48
Personnel research, 33, 38
  merit-rating, 206
  methods, 5–6
Personnel specifications, methods of determining:
  job analysis, 42–48
  judgments, 41–42
  statistical analysis, 48–51
Personnel Test, Form D., 113n
Personnel tests, 39, 51, 73–74, 76
  combining, 105–108
  extent of use, 76–77
  factors affecting efficiency of, 99–105
  functional utility, 91–105
  principles of, 75–109
  purposes, 75–76
  reliability, 92–96
  selection ratio, 96–99
  testers, training of, 108–109
  types, 79–81
  validity, 78, 91–92, 94–96
  (*see also* Employment tests)
Peters, C. C., 498n
Peterson, A. P. G., 451
Peterson, J., 23, 23n
Phelan, R. F., 179, 179n, 181
Philadelphia Quartermaster Depot, 193, 193n
Phorias, 147–148
Phrenology, 2, 3
Physical effort, 41–42
Physically handicapped (*see* Disabled)
Physician, 146
Physiognomy, 2, 3
Physiological fatigue, 434, 436
Piece work payment, 281–283, 337 (*see also* Wages)
Piersal, D. T., 347n, 350n, 351
Pitney, R. W., 53, 53n
Placement, 39
  bottle selectors, 388
  efficiency, 89, 97, 101
  electrical inspectors, 387–388
  employees to reduce accidents, 427–428
  function of tests, 77–78
  inspectors, 384–388
  methods of, 2–3
  program, 11–12
  use of tests, 75–76
  work performance, 12
Planty, Earl G., 296
Point method, 495–496
Polygon, 514
Pond, Millicent, 12, 12n, 116, 116n, 122, 122n
Pond, S. A., 297, 297n
Positive incentive, 337
Potter, C. J., 429, 429n
Power-sewing-machine operators, 130
Power test, 93
Practice distribution, 271
Praise, 268
Precision instruments, accuracy in use of, 380–384
Predictive validity, 92, 94
Predictors, 62–65
Price, Dennis H., 197n
Pride, in product, 313
Probability, 48
Probationary period, 3
Proctor & Gamble, 8
Production:
  illumination, 326
  morale, 304
  music, 455
  relation to work hours, 455
  rest pauses, 457
  training, 276–280
  training and reduction of, 280–281

Production *(cont.)*
  training and standards of, 281, 282 (*Fig.*)
  unit, 25, 30
Production efficiency, 279
Production records, 36
Productivity, 32
  attitudes, 333
  group behavior, 368–369
  hosiery loopers, 25–26
  individual differences, 24–27
  predictions, 28–29
Product-moment coefficient, 537–547
Projective tests, 163
Promotions, 32, 34–35, 268, 337
  achievement tests, 199–200
  merit-rating, 204–250
  personality traits, 162
  use of tests, 75–76
Prudential Life Insurance Company, 328, 341
Psychological Corporation, 558
Psychological fatigue, 436
Psychological Institute, 558
Psychological methods:
  acceptance of, 7–9, 7 (*Table*)
  personnel tests' development, 74
Psychological tests, 161
  fallibility, 78
  industrial skepticism, 76
Psychologists, 1, 6, 89, 108, 110, 269
  experiments by, 3
  individual differences, 30
  job attitudes, 161
  predictions, 98
  personality traits, 161
  wage factors and, 486
Psychology, 166
  applied, 8
  educational, 261, 265
  industrial, 1, 2
    application of, 6–7
  motion and time study, 475 475n
  use of, 294
Psychomotor abilities, 110
Psychomotor coordination 135
Psychomotor factors, 134–135
Psychomotor spread, 134
Psychomotor tests:
  correlation studies, 141–145
  definition, 134
Psychomotor tests *(cont.)*
  inter-correlation, 135
Psychopathic deviate scale, 169
Psychotechnology, 475n
Pullman Company, 430
Punch press test, 80
Purdue Achievement Tests, 79
Purdue Clerical Adaptability Test, 90, 156, 194
Purdue Grooved Pegboard Test of Manual Dexterity, 79, 107, 138, 138 (*Fig.*), 140 (*Table*), 142 (*Table*)
Purdue Hand Precision Test, 105, 385, 386
Purdue Mechanical Adaptability Test, 79, 126, 129, 130, 130 (*Fig.*), 132
Purdue Non-Language Personnel Test, 114
Purdue University, 152, 257, 258
Purdue Vocational Tests, 16, 195–199

Q

Qualification tests, 294
Questionnaires, 304
  job satisfaction, 319
  survey, 318 (*Table*)
  value of, 72–73
  (*see also* Opinion survey)
Quotas, 368–369

R

Radiation, 445–449
Radio Corporation of America:
  Camden works, 407
  employee preference for music, 455
Railway motormen, 419
Rand, G., 150n
Raney, E. T., 359 359n
Ranking method, 493
Rank-order system, 213
Raters, influences on, 228–233
Rating, 35, 61–62, 83 (*see also* Merit-rating)
Rating scale, 35 (*Fig.*)
Raube, S. A., 323, 323n
Real incentive, 337–338
Real wages, 485
Recognition training, 274
Records, 83
Recreation, 171

References, 39
  letters of recommendation, 71–72
  questionnaires, 72–73
Refrigerator-plant mechanics, 130
Regression, 338
The Relations of Hearing Loss to Noise Exposure, 452n, 453n
Relative validity, 158–160
Reliability, 91, 93–96, 128–129
  judgments, 342
  merit-rating, 242–244
  O'Connor Wiggly Block Test, 129
  and validity, relationship between, 94–96
Reliability coefficient, 126
Remmers, H. H., 128, 128n
Renshaw, S., 274, 274n
Rest pauses, 441
  efficiency and, 457–458
  length, 458
  nourishment, 459
  systematizing, 459
Resignation, 339
Retirement, 2 (*see also* Pension plans)
Revised Minnesota Paper Form Board, 127, 127 (*Fig.*)
Reymert, M. L., 242. 242n
Rhathymia, 167
Rice, S. A., 52n
Richardson, H. M., 165, 165n
Roach, D. E., 345, 345n
Roberts, E. B., 323, 323n
Roethlisberger, F. J., 368, 368n
Role playing, 295
Roofer, 187–188
Rorschach Projective Test, 170–171
Rothe, H. F., 233, 233n, 319, 319n
Ruch, F., 130n
Rundquist, E. A., 36, 36n, 245, 245n
Rusmore, J. T., 177, 177n, 178
Russ, J. J., 323, 323n, 325
Russell, J. T., 96n, 102n
Ruter, W. H., 331, 331n, 332, 332n
Ryan, T. A., 438, 438n, 439

S

Safety, 27, 398–432 (*see also* Accident proneness *and* Accidents)
Safety conferences, 428–429
Safety director, 400, 402, 418–419
Safety program, 6, 427–432
Salary, 206 (*see also* Wages)
Sales checking test, 117
Salesclerks, 50–51, 171
Sampling statistics, 513
Satter, G. A., 502n
Scheidt, V. P., 450, 450n
Schell, J. W., 128, 128n
Schizophrenia scale, 169
Schmidt, H. O., 169, 169n
Schmitt, E., 416, 416n
Schutz, H. G., 509, 509n
The Science of Seeing, 444n
Science Research Associates, 558
Scientific development department, 172
Scientific management, 294–295
Screw manufacturing company, 132
S. D. (*see* Standard deviation)
Seashore, R. H., 134, 134n, 142 (*Table*), 290, 290n, 291
Seashore, H. G., 195, 195n
Secretaries, 240–241
Security, 313
Selection:
  achievement tests, 199
  apprentices, 294
  function of tests, 77–78
  of inspectors, 384–388, 397
  of supervisors, 343
  test methods, 192–193
  use of tests, 75–76
Selection procedure, flow chart, 40 (*Fig.*)
Selection program, work performance, 12
Selection ratio, 85, 91, 96–99, 101
Selover, R. B., 77, 77n
Semanek, Irene, 126n, 131n
Senders, Virginia L., 465, 465n
Seniority, 3, 13, 162, 208–209, 209 (*Table*)
Sewing machine operators, 363–364, 365 (*Fig.*)
Shartle, C. L., 38n, 45, 45n, 174, 174n, 186, 186n, 503n
Sheridan Supply Company, 558
Ship recognition training, 274
Shipyard machinist, 189 (*Table*)
Shrosbee, G., 417, 417n
Schuman, J. T., 132, 132n
Sight-screener, 150, 150n
Sisson, E. D., 219n
Sitzungsbericht der Akademie der Wissenshaft, 442n
Sleight, R. B., 442, 442n, 463, 463n

Slocombe, C. S., 403, 403n
Smith, H. C., 336, 336n 455, 455n
Smith, M., 103, 103n, 476n
Smyth, R. C., 495, 495n, 503n
Snell, Albert C., 148n, 149n
Snellen letter chart, 148, 148 (*Fig.*)
Snellen notation of acuity, 148–149
Social introversion-extroversion, 167, 169
Social needs, 336
Sound (*see* Noise)
Space test, 115
Spatial ability, 158
Spatial relations, 124
Spatial visualization, 111
Spears, E. M., 454, 454n
Specialized aptitudes, 110
Speed test, 93
Speroff, B., 426, 426n
Split-halves method, 93
Squires, P. C., 472, 472n, 473
SRA Primary Mental Abilities Tests, 114–116
Stability, job performance, 28–29
Stagner, R., 4, 4n, 7n
Standard deviation, 21, 21n, 525–527, 547
Standardized interview, 54
Standard typing test, 80
Stanford Motor Skills Unit Test, 142 (*Table*)
Stanford University Press, 558
Stapleton, E. S., 177, 177n, 178
Starling, E. H., 434, 434n
Starr, R. B., 210n, 211n
Statistical analysis, 62
  application form development, 67–70
Statistical process, factor analysis in, 92
Statistical comparison, factors of, 78
Statistics:
  arithmetic mean, 519–521, 524
  central tendency measures, 519–524
  comparable scores, 527–532
  correlation, 534–547
  descriptive vs. sampling, 513–514
  frequency distribution, 514–518
  graphic representation of data, 514–519
  mean, 547
  median, 521–524
  mode, 523–524
  normal distribution, 518
  percentiles, 532–534
Statistics (*cont.*)
  polygon, 514
  product-moment coefficient, 537–547
  standard deviation, 525–527, 547
  variability, 524–527
Stead, W. H., 38n, 186, 186n, 503n
Steel mill employees, injuries of, 430–431
Steinberg, M. D., 90, 90n, 91 (*Fig.*)
Stenographers, 174
Stenographic tests, list of, 555
Stereopsis, 146–147
Sterling, Scott, 149n
Sterner, J. H., 453, 453n
Stetson, R. H., 477n
Stevason, C. C., 16n, 196n
Stevens, A. F., Jr., 417, 417n
Stevens, S. N., 117, 117n
Stevens, S. S., 450, 450n
Steward, Verne, 64, 64n
Stimulus, 335
Stock, F. G. L., 322, 322n, 325, 449, 449n, 455, 455n
Stockford, L., 89n
Stoelting, C. H., Company, 558
Stone, H. C., 62, 62n
Stromberg Dexterity Test, 139
Strong, E. K., Jr., 171, 171n
Strong Vocational Interest Blank, 171–172, 174
Structured interview, 54
Subjective fatigue, 436, 437 (*Table*)
Substitute incentive, 337–338
Supervisors, 27, 41, 84, 204, 375, 421
  absenteeism, 355, 356 (*Table*)
  adaptability test scores and job success, 119 (*Fig.*)
  aid to job analysis, 490–491
  attitudes, 341–343
  behavior, 345–346
  conferences, 285–286
  disputes, 4
  employees, relations with, 301, 304, 341–342, 344
  group unity, 371
  human relations, 340–346
  influence on job satisfaction, 344
  interest test, 174
  mental ability tests, 118–120
  merit-rating, 204, 206–207
  powers, 3
  rating tendency, 222
  selection of, 343

Supervisors (*cont.*)
  self-improvement, 344–345
  training by, 1, 4–5, 251, 270–271, 289
  training check list, 255 (*Table*)
  training of, 4, 273, 288. 294–301, 343
  wage rates, 487–488
Supply and demand, relation to wages, 483–484, 488
Surgent, L. V., 104, 104n
Survey Research Center, 328, 331n, 341n
  University of Michigan, 345n
Synthetic test validation, 89–90, 92

T

Tardiness, 329, 332
Taylor, C. W., 201, 201n
Taylor, E. K., 38, 38n
Taylor, H. C., 96n, 102n, 243n
Taylor, H. R., 276, 276n
Taylor-Russell tables, 102, 108, 109, 548–553
  analysis of, 103–105
  use in predicting job success, 102–103
Team spirit, 303 (*see also* Morale)
Technical training, 288 (*see also* Training)
Technical information test for electricians, 197
Technical Information Test for Machinists and Machine Operators, 16, 196
Telebinocular, 150, 150n
Telephone repair crew, 365–367, 367 (*Fig.*)
Teletype operator trainees, 120–121
Temperament:
  fluctuations in, 421
  paranoid trait, 167–168
  traits, 163
Temperature, 445–449
Template, 70
Tenure, 32, 34, 57 (*see also* Employment stability)
Termination, 9, 316–317
  factors preventing, 13–14
  merit-ratings, 206
  (*see also* Firing *and* Layoffs)
Terrell, C. M., 274, 274n
Test administrator, 108–109
Testing, principles of, 75–109
Testing program:
  averages, 78
  validated tests, use of, 81–82
Test procedures, standardization of, 32
Test reliability, importance of, 93–94
Test-retest method, 92–93
Tests:
  achievement, 79, 86, 185–203
  adaptability, 79, 113–114, 117, 121, 126
  adult placement, 120
  aptitude, 79–81, 110–160, 294
  battery of, 105–108
  blueprint reading, 197
  change making, 117
  for children, 113, 115
  clerical, 116, 156, 194–195
    list of, 555
  completion form, 297
  dexterity, 32, 93, 140–144
    list of, 556
  dotting, 140 (*Fig.*)
  economic conditions and, 77–78
  electricians, 197
  employee, 77
  employment, 40–41, 53n, 76, 81
  engine lathe operation, 198
  forced-choice, 176–178
  fork-lift operators, 193–194, 193 (*Fig.*)
  functions of, 77–78
  group, 79, 112
  individual, 79
  industrial mathematics, 197–198
  industrial training classification, 126
  inspection, 157
  instrumental, 79
  intelligence, list of, 554–555
  interchangeability, 80–81
  interest, 79, 171–174
    list of, 556
  inventories, 79
  job miniature, 190
  job sample, 375–385
  lists of, 554–557
  machine shop, 196
  manipulative, list of, 556
  mechanical, list of, 556
  mental ability, 111–122
  motor, 115
  paper and pencil, 79, 112, 139, 163
  perception, 115, 155–156
  personality, 79, 81, 87, 163–171
  personnel, 39, 51, 73, 109
  placement, 77–78
  power, 93

- Tests (*cont.*)
  - psychological, 76
  - psychomotor, 134–145
  - publishers of, 557–558
  - punch press, 80
  - sales checking, 117
  - selection, 77–78
  - space, 115
  - speed, 93
  - standard statistical, 48
  - testing of (*see* Test validation)
  - typing, 80
  - use by large companies, 76
  - vision, 32, 145, 148–156
    - list of, 555
  - (*see also under name of test*)
- Test validation:
  - administering, 82–83
  - analysis of result, 84
  - coefficient of correlation, 84–85
  - correlation with ability, 86
  - correlation with experience, 87
  - criterion, 83
  - employee anxiety, 82–83
  - follow-up method, 85
    - critique 86–88
  - job-group method, 89
  - labor organizations, 83
  - management, 83
  - methods, 82–91
  - present-employee method, 82–85
    - critique, 86
  - synthetic validity method, 90
  - test scores and job success, 121, 121 (*Table*)
- Thinking introversion-extroversion, 167
- Theme analysis, 313–315
- Thorndike, E. L., 222, 222n, 268n, 270, 270n, 271n, 304
- Thornton, G. R., 126n
- *Thurstone Employment Tests, Examination in Typing*, 194n
- *Thurstone Examination in Typing*, 80, 194–195, 195 (*Fig.*)
- Thurstone, L. L., 111, 111n, 123, 123n, 124, 124n, 157n, 169n, 172, 172n, 217, 217n, 305, 305n
- *Thurstone Temperament Schedule*, 169–170, 309–310, 319, 437
- Tiffin, Joseph, 2, 2n, 13, 48, 48n, 113n, 114n, 117, 117n, 122, 122n, 131n, 136n, 138n, 141, 141n, 151n, 179, 179n, 181, 190, 190n, 198n, 239n,
- Tiffin, Joseph (*cont.*) 271n, 359, 359n, 380, 380n, 383, 383n, 414n, 482, 482n, 506, 506n
- Time study, 45, 276 (*see also* Motion and time study)
- Tinker, M. A., 444, 444n
- Tin-plate inspection:
  - accuracy of, 374–380
  - methods of, 389–391
- Tin-plate inspectors:
  - placement of 384–387
  - training of, 395–397
- Tiredness, 439
- Toedt, Theodore A., 5n
- Topal, J. R., 458, 458n
- Touch type system, 290
- Trabue, M. R., 139n, 142 (*Table*)
- Tracing test, 140 (*Fig.*)
- Trade school, 127
- Trade tests, list of, 557
- *Trade Test Standard Manual for Written and Practical Tests*, 500
- Training, 24, 39, 44, 58–59, 79, 127, 251–301 492
  - absenteeism reduction, 284–285
  - accident reduction, 284
  - aids, 272–287
  - apprentices, 294, 396–397
  - breakage, decrease in, 283
  - cost, 14, 33–34
  - employee, 2–3
  - experience, effect on, 277–279
  - experts, use of, 288
  - individual differences, 20–24, 31–32
  - in-service, 286
  - instructors, 263, 267, 289, 291–292
    - Air Force, 220
  - interviewers, guide for, 58–59 (*Table*)
  - leadership, 300–301
  - learning:
    - characteristics of, 261–265
    - principles of, 265–272
  - management, 294–301
  - measurement method, 261 (*Fig.*), 296–297
  - methods, 267, 272–287, 295–296, 396
  - supervisors, 4–5, 294–301, 343
  - tin-plate inspectors, 395–397
  - transferability, 269–271
  - turnover reduction, 285
  - use of tests, 75–76, 200–202
  - value of tests, 133
  - (*see also* Learning)

Training needs, 295
  check lists, 253, 254 (*Table*), 255 (*Table*)
  determination of, 252–261, 253 (*Fig.*)
  job analysis, 256–261
  judgments, 254–256
  manual, 258
Training programs:
  cost reductions, 280–281
  production standards, 281
  purposes, 251–252
  types, 287–301
Transactions of the American Academy of Ophthalmology and Otolaryngology, 444n
Transfers:
  achievement tests, 199–200
  merit-rating, 204–250
  use of tests, 75–76
Troland, L. T., 442, 442n
True-false tests, 188 (*Table*), 297–299
Tufts College Institute for Applied Experimental Psychology, 462n
Turnover, 48, 331
  cost of, 3–4
  reduction through statistical analysis, 49–50
  test scores, 116–117
  (*see also* Labor turnover)
Typing tests, list of, 555

U

Unguided interview, 315–316
Uhrbrock, R. S., 40, 40n, 40 (*Fig.*), 217, 217n, 219, 308, 308n
Uhrbrock attitude scale, 309 (*Table*)
Union contract, 286
  merit-rating, 208–209
  readability, 358–360
  wages, 489
Unions:
  attitude toward merit-ratings, 208–209
  output restrictions, 369
  recognition of, 4–5
  seniority, 208–209
  type of wage demands, 483 (*Table*)
  wage demands, 482, 500
  (*see also* Labor *and* Organized Labor)
United States, 475n
United States Department of Commerce, 53n, 228n
United States Department of Labor, Employment Service, 44n, 186
United States Rubber Company, job analysis chart, 259–260 (*Table*)
United States War Department, 193n
University Book Store (Purdue), 558
Upgrading program, 500
Upjohn Foundation for Community Research, 243n
U.S.A.F.:
  aviation psychology program, 469n
  Wright Air Development Center, 465n

V

Validation:
  of aptitude tests, 157–160
  of oral trade questions, 186–187
  of personality and interest test, 179–181
Validity, 91–92, 94–95, 98, 101, 128
  of achievement tests, 185–186
  concurrent, 94
  of judgments, 60–61
  of merit-rating, 244
  of miniature punch press test, 190–192
  predictive, 94
  and reliability, relationship between, 94–96
  of tests, 91–92
Validity coefficient, 96 (*Fig.*), 98–99, 106
Values, 337
  employees, 320 (*Table*), 321 (*Table*), 322–326
Value system, 320
VanderMeer, A. W., 272n
Van Voorhis, W. R., 498n
Van Zelst, R. H., 330, 330n, 371, 371n, 416, 416n
Variability, 22, 26, 27, 88
  of ratings, 238–240
Variability measures, 524–527
Variables, 6, 36, 327, 329
Verbal comprehension, 111
Vernon, H. M., 409, 409n, 411, 411n, 437, 437n, 448, 448n, 457, 457n
Vernon, P. E., 3, 3n, 135, 135n
Versatility, 32
Vestibule training, 288

Victor Adding Machine Company of Chicago, 310
  opinion survey chart, 311 (*Fig.*)
  questionnaire, 312
Vision, 42, 80
  accident causative, 413–419
  color, 147
  filming of eye movement, 394–395
  foveal, 394
  illumination, 441–442
  job success, 145–146
  peripheral, 394
  space visualization, 42
  visual displays, 463
Vision tests, 32, 145, 148–149
  devices, 150–151
  list of, 555
  personnel placement, 146–156
  relation to accidents, 427
  stereoscopic method, 150–151
Visual acuity, 146, 148–149, 441–442
  conversion scale, 149 (*Fig.*)
  illumination, 441–442
  tests, 385
Visual displays, 463–465
  use of dials, 464 (*Figs.*)
Visual efficiency, 149
Visualization, 123–124
Visual job families, 152–153
  job success and visual skill profiles, 155 (*Fig.*)
  skill test requirements, 154 (*Fig.*)
Visual performance, 146–148
Visual profile, 152
Visual skills, 110, 145–156
  classification by job families, 152–156
  jobs requiring, 145–146
Viteles, Morris S., 23, 23n, 360, 360n, 368, 368n, 403, 403n, 458, 458n
Vocational guidance, 171–172, 183–184
  influence on test answers, 174
Vocational schools, achievement tests, 202–203
Vocational teachers, 195
Vogel, J., 77, 77n

W

Wadsworth, G. W., 120, 120n, 165n
Wage payment plans, 267–268, 292
Wages, 313, 321–323, 482–512
  collective bargaining agreement, 489
Wages (*cont.*)
  company civic relations, 484
  company policy, 487–488
  cost of living, 485
  discrimination, 487
  hourly paid job, 486–487
  incentive plans, 24, 457, 486–487
  industrial differentials, 485
  job analysis, 489
  merit-rating, 206, 500
  minimum, 14, 34
    legislation, 107, 484
  organized labor, 485–486
  piece work, 281–283
  regional differentials, 485
  seniority, 500
  supervisors, 487–488
  supply and demand, 483–484
Wage Stabilization Board, 484
Wagner Act (*see* National Labor Relations Act)
Wagner, R., 60, 61n
Wagner, T., 47, 47n
Walker, W. S., 429, 429n
Walsh, F. X., 359, 359n
Wantman, M. J., 53, 53n
Ware, H. F., 285, 285n
War Manpower Commission Bureau of Training, 294n
Warner, C. G., 448, 448n
Waters, J. F., 316–317, 316n
Weighted check list, 216–219, 218 (*Tables*), 219 (*Table*)
Weitz, J., 331, 331n
Welders, 132
Welding test, 197
Wells, David 151, 151n
Wells, R. G., 126n
Western Electric Company, 135, 135n, 326, 422
  Hawthorne plant, 368
Weston, H. C., 450, 450n
Wetzel, M., 442, 442n
Wheeler, D. E., 452, 452n, 453
Wherry-Doolittle Technique, 503n
Whitehead, T. N., 326
White Motor Company, 58
Wilensky, H., 329, 329n, 367, 367n, 368, 368n
Wilson, R. F., 503n, 509, 509n
Winn, A., 509, 509n
Winslow, C. E. A., 446, 446n, 447

Winstanley, N. B., 188, 188n, 189n
Wirt, S. E., 151n
Women, wages of, 487
Wonderlic, E. F., 56, 56n, 113, 117, 117n
Wonderlic Personnel Test, 79, 113–114, 126
Wong, W. A., 400, 400n
Wood, W. F., 254, 254n, 256
Woods, H. M., 403, 403n
Woodson, W. E., 462n
Woodworth, R. S., 270, 270n
Woodyard, E., 271n
Word fluency, 111
Work area layout, 394
Work-group unity, 368–370
Work history, value of, 62–65
Work hours, 455–457
Working conditions, 433–459
  atmospheric conditions, 445–449
  boredom, 436
  efficiency factor, 438–439
  fatigue, 434, 436
  Lauru platform, 438
  hearing loss, 453
  hours of work, 455–457
  illumination, 441–445
  impairment, 436
  music, 454–455
  noise, 450–454
  output factor, 437–438
  physiological factors, 434–436
Working conditions (*cont.*)
  psychological factors, 436–437
  relationships among criteria, 439–440
  rest pauses, 457–459
Workmen's compensation, 146
Work methods:
  and human engineering, 460–481
  (*see also* Human engineering)
Work output, relation to working conditions, 437–438
Work sample test, 190
World Book Company, 558
World War II, 274, 460, 484
Worry, 419
Worthy, J. C., 334, 334n, 371, 371n
Wright, H. A., 422
Wright Air Development Center, 465n
Written achievement tests, 195–199
Wundt, Wilhelm, 1
Wyatt, S., 322, 322n, 325, 449, 449n, 455, 455n

Y

Yoder, Dale, 484, 484n, 486, 486n
York Knitting Mills, 58

Z

Zalkind, S. S., 253, 253n, 254, 274, 274n, 294, 294n
Zerga, J. E., 45, 45n, 210n

BUREAU OF EDNL. & PSYL. RESEARCH

C

Form No. 3.

PSY, RES.L-1

Bureau of Educational & Psychological Research Library.

The book is to be returned within the date stamped last.

29.5.60		
30.9.67		

WBGP-59/60-5119C-5M

658
TIF
Form No. 4

BOOK CARD

Coll. No. 658 Accn. No. 1419
Author. McCORMICK ERNEST J.
Title INDUSTRIAL psychology.

Date.	Issued to	Returned on
29.5.60	J.C.D.	Rtd. [illegible]